



Universiteit
Leiden
The Netherlands

Improving management of infectious diseases at the emergency department: from risk prediction to treatment strategies

Kaal, A.G.

Citation

Kaal, A. G. (2026, September 16). *Improving management of infectious diseases at the emergency department: from risk prediction to treatment strategies*. Retrieved from <https://hdl.handle.net/1887/4309760>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309760>

Note: To cite this publication please use the final published version (if applicable).

Improving management of infectious diseases at the emergency department

From risk prediction to treatment strategies



Anna G. Kaal

Improving management of infectious diseases at the emergency department

From risk prediction to treatment strategies

Anna G. Kaal

Improving management of infectious diseases at the emergency department: from risk prediction to treatment strategies

© 2026, Anna Kaal, The Netherlands

Cover design: Lotte Hoeksma and Anna Kaal

Layout: Arina van Londen | Ridderprint, www.ridderprint.nl

Printed by: Ridderprint

Financial support for the printing of this thesis: Wetenschapsfonds HagaZiekenhuis and department of Biomedical Data Sciences LUMC

ISBN: 978-94-6537-594-6

Improving management of infectious diseases

at the emergency department:

from risk prediction to treatment strategies

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden

op gezag van rector magnificus prof.dr. S. de Rijcke,

volgens besluit van het college voor promoties

te verdedigen op woensdag 16 september 2026

klokke 16:00 uur

door

Anna Gonda Kaal

geboren te Amsterdam

in 1992

Promotor:

Prof. dr. E.W. Steyerberg

Copromotor:

Dr. C. van Nieuwkoop

Promotiecommissie:

Prof. dr. S. le Cessie

Prof. dr. K.D.P.W.B. Nanayakkara, Amsterdam UMC

Prof. dr. M.J.M. Bonten, UMC Utrecht

Prof. dr. L.G. Visser

TABLE OF CONTENTS

Chapter 1	General introduction	7
Part I - Diagnostics: Improving the diagnostic process of bacteremia using procalcitonin		
Chapter 2	The diagnostic accuracy of procalcitonin for community-acquired bacteraemia: an updated systematic review and meta-analysis	17
Chapter 3	Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model	47
Chapter 4	Bacteremia prediction models to reduce unnecessary blood cultures: external validation in a large US emergency department	73
Part 2 - Risk stratification in patients with infectious diseases at the emergency department		
Chapter 5	Analysis of variation between diagnosis at admission vs discharge and clinical outcomes among adults with possible bacteremia	97
Chapter 6	Outcomes of community-acquired pneumonia using the Pneumonia Severity Index versus the CURB-65 in routine practice of emergency departments	119
Chapter 7	Diagnostic yield of bacteriological tests and predictors of severe outcome in adult patients with COVID-19 presenting to the emergency department	137
Part 3 - Treatment: Oral treatment of infectious diseases at the emergency department		
Chapter 8	Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study	159
Chapter 9	Choose oral over intravenous antibiotic therapy	181
Chapter 10	General discussion	191
Appendices	Nederlandse samenvatting	212
	Dankwoord	215
	List of publications	217
	Knowledge dissemination	218
	Curriculum Vitae	219



Chapter 1

General introduction

GENERAL INTRODUCTION

Infectious diseases remain a leading cause of morbidity, mortality and associated healthcare utilization.¹ This burden is increasing, due to various factors such as population growth, aging populations, climate change and increasing antimicrobial resistance.^{2,3} Emergency departments (EDs) play an essential role in the recognition and management of infections. Patients present with a broad spectrum of symptoms, ranging from a simple cough caused by a self-limiting respiratory tract infection, to life-threatening septic shock requiring rapid intensive care treatment.⁴ Unfortunately, although in some patients symptoms are evident, many symptoms are relatively unspecific and recognizing those with a high risk of a severe disease course is challenging.

Three important steps in the management of infectious diseases are ordering the right diagnostic tests, assessing the risk of complications and deciding on appropriate treatment. By ordering the right tests, one aims to make an accurate diagnosis and prognosis, which is essential to decide on targeted treatment. However, there are downsides to exposing more patients than needed to testing and treatment. For instance, false-positive blood cultures are associated with increased length of stay, increased exposure to broad-spectrum antibiotics, increased costs and even mortality.⁵⁻⁷ Risk stratification can help in deciding on site of care and intensity of monitoring or treatment. The subsequent treatment strategy is aimed at the presumed causative pathogen, ideally with the most precise effect possible. Overtreatment with antibiotics can lead to unnecessary hospital admissions, therapy-related adverse events, and antimicrobial resistance.⁸ Therefore, clinical decision-making often involves balancing the risks of doing too little against the risks of doing too much.

Finding this balance can improve individual patient outcomes, but can also play a role in addressing broader healthcare challenges. Many countries face increasing pressure on their healthcare systems, driven by rising demand and shortages of resources such as personnel and materials.^{9,10} Identifying those patients who actually benefit from diagnostic testing and more intense treatment strategies may help cope with these problems by reducing workload and reducing hospital admissions and associated costs. Furthermore, overuse of diagnostic tests and exposure to unnecessary antibiotics drive antimicrobial resistance.^{3,11,12} Antimicrobial resistance is a serious health threat, and was associated with 36% of global infection-related deaths in 2019.¹² This number is expected to increase substantially by 2050 if left unaddressed.¹² With this in mind, antimicrobial stewardship programs have been widely implemented over the last decades to optimize antibiotic use.⁸ More recently, the contribution of false-positive tests to antimicrobial resistance has initiated a call

for 'diagnostic stewardship'. This initiative aims to order the right test, at the right moment for the right patient.¹¹

OVERALL AIM

This motivates the aim of this thesis: to identify and evaluate strategies to optimize the management of infectious diseases in adult ED patients, with the objective of ensuring that diagnostic testing and treatment are provided to patients who require them, while minimizing unnecessary interventions. We consider three different aspects relevant for management of infections: diagnostics, risk stratification and treatment. This thesis focuses on two common and clinically important infectious diseases: bacteremia and community-acquired pneumonia.



Figure 1. Outline of this thesis with overall and specific aims.

OUTLINE OF THESIS AND SPECIFIC AIMS

The first part of this thesis focuses on improving the *diagnostic* process of bacteremia, aiming to safely reduce blood cultures in patients with suspected community-acquired bacteremia. This could improve patient outcomes by preventing negative consequences of false-positive blood cultures, while decreasing costs and resources utilization.

Chapter 2 presents a systematic review and meta-analysis of the diagnostic accuracy of the biomarker procalcitonin (PCT) for bacteremia.¹³ Procalcitonin is a biomarker often used to differentiate between a viral versus a bacterial infection.¹³ Under normal conditions, PCT is produced in negligible amounts by thyroid C-cells. In case of a bacterial infection, levels rise rapidly in response to pro-inflammatory

cytokines and bacterial toxins. Importantly, viral infections suppress PCT production through interferon- γ .¹⁴ This makes it a relatively specific marker for bacterial infections, in contrast to commonly used biomarkers such as C-reactive protein or white blood cell count. We investigate whether low values of PCT can help rule out bacteremia, and hence reduce unnecessary blood cultures.

Chapter 3 and **Chapter 4** show the development and validation of a diagnostic prediction model estimating the risk of bacteremia.^{15,16} When trying to identify which patients benefit from testing or treatment, one approach is to use predictive modeling. Prediction models are divided in *diagnostic* and *prognostic* models (**Table 1**).¹⁷ Diagnostic models aim to estimate the probability that an individual patient currently has a specific condition. This information can be used to support decisions about further testing or treatment. For the context of infectious diseases, the Shapiro score is an example of a diagnostic model.¹⁸ This model estimates whether the patient has a high or low risk of bacteremia, before blood culture results are available. However, this score has not found its way to clinical practice. We hypothesize that this complex model is too labor-intensive for busy EDs, while simpler models lack diagnostic accuracy.¹⁹ Therefore, we aimed to develop and validate an easy-to-use model for bacteremia with high diagnostic accuracy. **Chapter 3** describes the development and validation of multiple prediction models that combine PCT with or without C-reactive protein and clinical characteristics.¹⁴ Based on lessons learned, we set out to update the most promising models and perform another external validation in Ann Arbor, United States.¹⁵ This is described in **Chapter 4**.

The second part of this thesis focuses on evaluation of clinical outcomes in a broad emergency department population with suspected infection, while exploring and evaluating *prognostic* models to identify patients at risk for a severe disease. This aims to improve risk-stratification to guide management strategies.

In **Chapter 5** we try to improve our understanding about misdiagnosis of infection, and whether this impacts clinical outcomes.²⁰ In addition, a prognostic prediction model for misdiagnosis was developed. As mentioned, prediction models are divided in *diagnostic* and *prognostic* models. Prognostic models aim to estimate the risk of a future clinical outcome, such as ICU admission or mortality (**Table 1**).¹⁷ This risk stratification can help to guide management strategies. **Chapter 6** and **Chapter 7** focus on patients with respiratory tract infections. First, **Chapter 6** answers the question whether the Pneumonia Severity Index (PSI) or the CURB-65 score, two of the most used prognostic prediction models for community-acquired pneumonia (CAP), are associated with better clinical outcomes in the Dutch population.²¹⁻²³ These models estimate the risk of 30-day mortality in CAP patients. The Dutch national guideline for CAP advises to use one of these models to decide on treatment

strategy.²⁴ Second, in **Chapter 7** we explore clinical outcomes and prognostic factors for a severe disease course in patients with COVID-19.²⁵

Table 1. Differences between diagnostic and prognostic prediction models.

	Diagnostic model	Prognostic model
Goal	Estimate current risk of having a specific disease	Estimate future risk of a clinical outcome
Example in this thesis	Shapiro score ¹⁸	PSI score ²² , CURB-65 ²³
Application	Support whether further testing or immediate treatment is needed	Guide decisions regarding level of care and treatment intensity

Finally, *the third part* of this thesis takes a different approach to improving the management of infectious diseases. Instead of identifying individual patients that benefit from specific interventions, the aim is to explore whether antibiotic *treatment* can be safely simplified from intravenous to oral treatment. This could prevent unnecessary exposure to broad-spectrum antibiotics and potentially reduce hospital admissions.

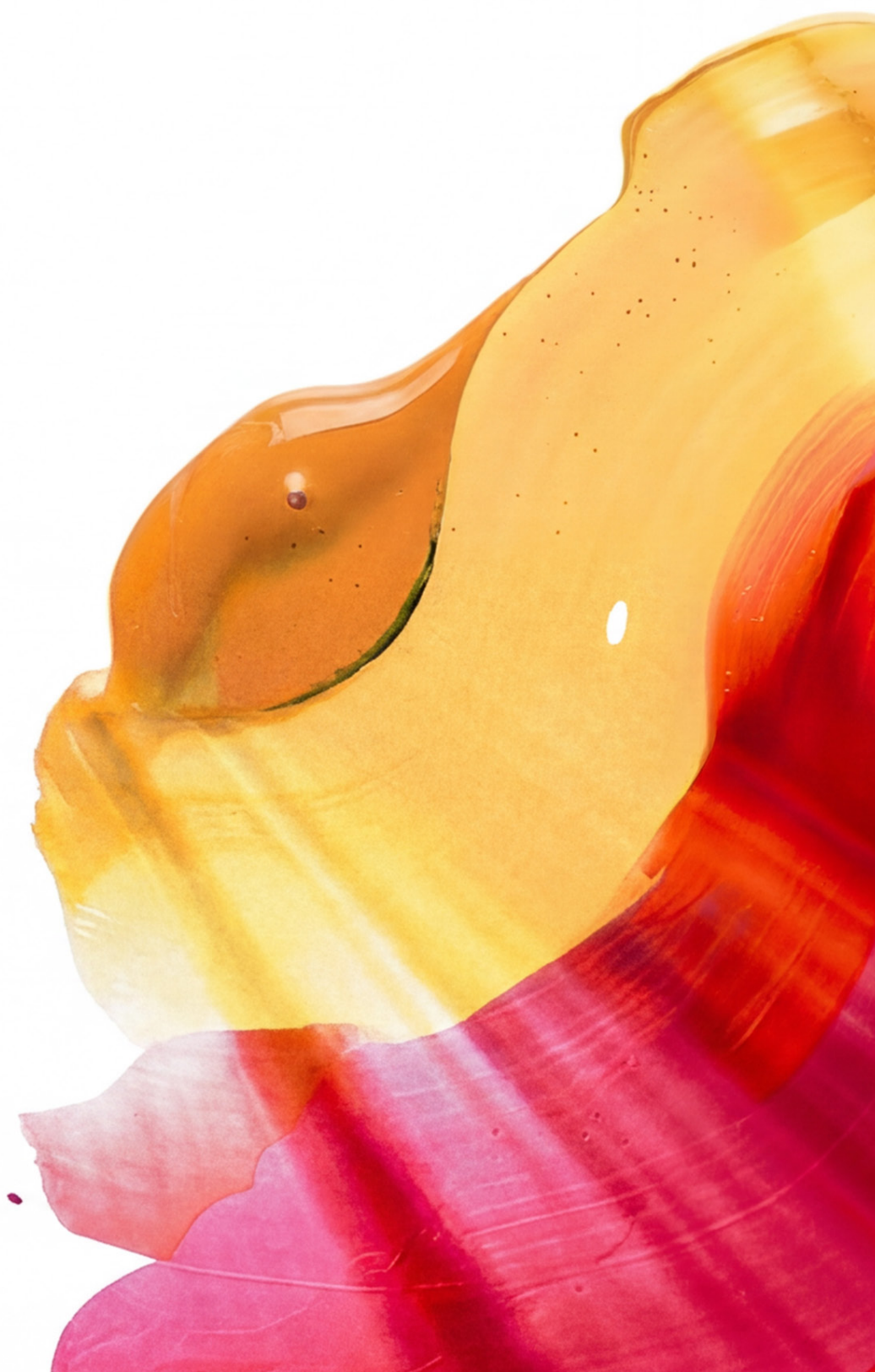
In **Chapter 8** we evaluate clinical outcomes for patients with moderate-to-severe CAP for those treated with oral antibiotics versus those treated intravenously.²⁶

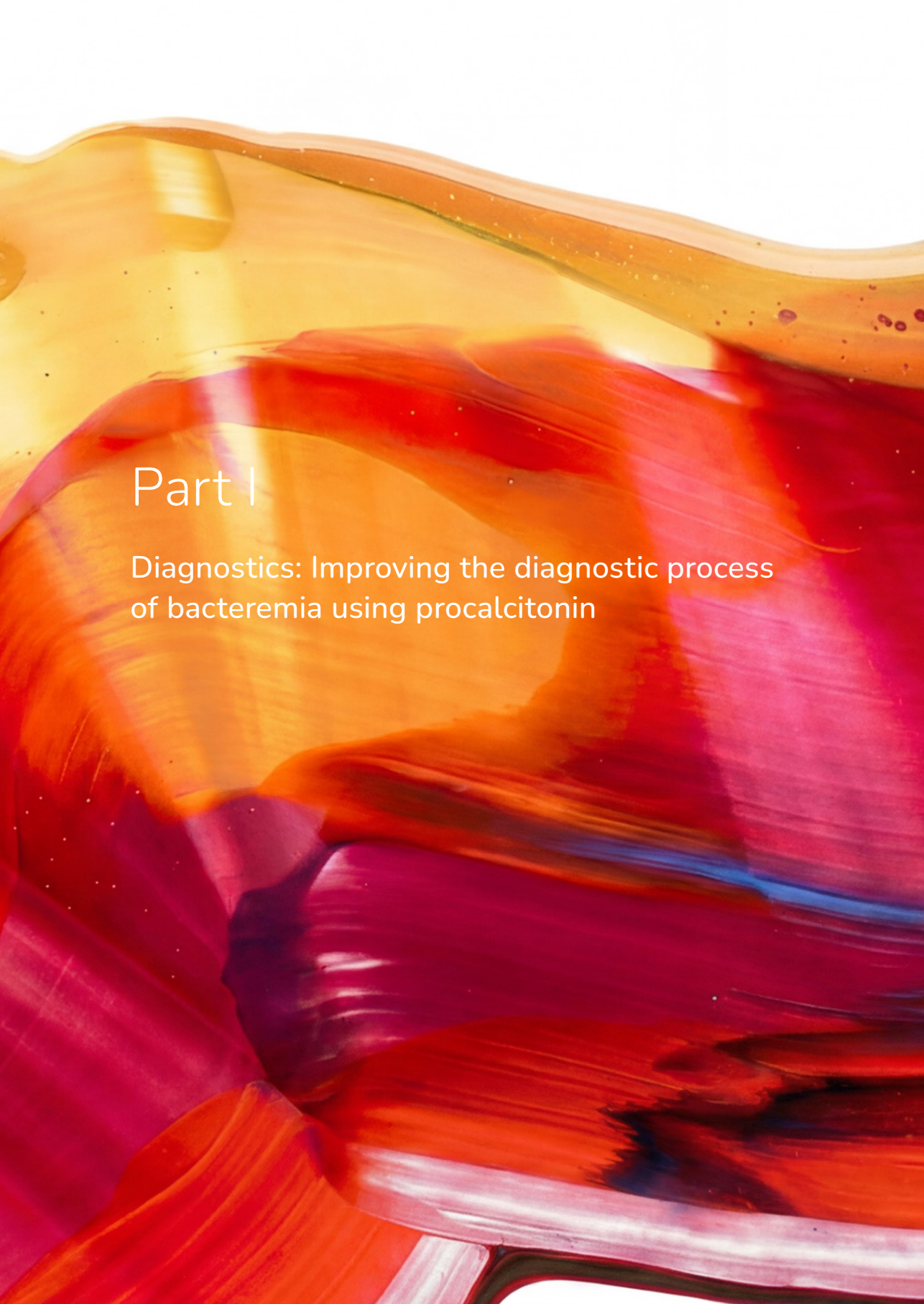
Chapter 9 is an opinion paper written for Dutch clinicians, explaining the context and rationale, advocating for a more prominent role of oral antibiotic therapy in a broader range of infectious diseases.²⁷ Finally, this thesis concludes with a general discussion, reflecting on the main findings, challenges, implications for clinical practice and directions for future research.

REFERENCES

1. WHO. Global Health Estimates: Life expectancy and leading causes of death and disability. <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates> accessed on 25th of June, 2025
2. Mora C, McKenzie T, Gaw IM, et al. Over half of known human pathogenic diseases can be aggravated by climate change. *Nat Clim Chang*. 2022;12(9):869-875. doi: 10.1038/s41558-022-01426-1. Epub 2022 Aug 8. PMID: 35968032; PMCID: PMC9362357.
3. GBD 2021 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet*. 2024 Sep 28;404(10459):1199-1226. doi: 10.1016/S0140-6736(24)01867-1. Epub 2024 Sep 16. PMID: 39299261; PMCID: PMC11718157.
4. Evans L. et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. 2021 Nov;47(11):1181-1247. doi: 10.1007/s00134-021-06506-y. Epub 2021 Oct 2. PMID: 34599691; PMCID: PMC8486643.
5. Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: importance, implications, interpretation and prevention. *Clin Microbiol Infect* 2018 Sep;24(9):964–9. <https://doi.org/10.1016/j.cmi.2018.03.030>. Epub 2018 Apr 3. PMID: 29621616.
6. Dempsey C, Skoglund E, Muldrew KL, Garey KW. Economic health care costs of blood culture contamination: a systematic review. *Am J Infect Control* 2019 Aug;47(8):963–7. <https://doi.org/10.1016/j.ajic.2018.12.020>. Epub 2019 Feb 20. PMID: 30795840.
7. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization. The true consequences of false-positive results. *JAMA* 1991 Jan 16;265(3):365–9.
8. Dyar OJ, Huttner B, Schouten J, Pulcini C; ESGAP (ESCMID Study Group for Antimicrobial stewardship). What is antimicrobial stewardship? *Clin Microbiol Infect*. 2017 Nov;23(11):793-798. doi: 10.1016/j.cmi.2017.08.026. Epub 2017 Sep 4. PMID: 28882725.
9. WHO. Global strategy on human resources for health: Workforce 2030. 2016
10. Centers for Disease Control and Prevention Health Alert Network. Disruptions in availability of Becton Dickinson (BD) BACTEC™ blood culture bottles blood culture bottles. 2024. <https://emergency.cdc.gov/han/2024/han00512.asp>. Accessed 25th of June 2025.
11. Zakhour J et al. ; International Society of Antimicrobial Chemotherapy (ISAC) and the Alliance for the Prudent Use of Antibiotics (APUA). Diagnostic stewardship in infectious diseases: a continuum of antimicrobial stewardship in the fight against antimicrobial resistance. *Int J Antimicrob Agents*. 2023 Jul;62(1):106816. doi: 10.1016/j.ijantimicag.2023.106816. Epub 2023 Apr 13. PMID: 37061101.
12. Ho CS et al. Antimicrobial resistance: a concise update. *Lancet Microbe*. 2025 Jan;6(1):100947. doi: 10.1016/j.lanmic.2024.07.010. Epub 2024 Sep 18. PMID: 39305919.
13. Kaal AG, Nieberg M, Stegmeijer K, Steyerberg EW, van Nieuwkoop C. The diagnostic accuracy of procalcitonin for community-acquired bacteremia; an updated systematic review and meta-analysis. [Submitted for publication]
14. Samsudin I, Vasikaran SD. Clinical Utility and Measurement of Procalcitonin. *Clin Biochem Rev*. 2017 Apr;38(2):59-68. PMID: 29332972; PMCID: PMC5759088.

15. Kaal AG, Meziyerh S, van Burgel N et al. Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model. *J Infect.* 2024 Oct;89(4):106251. doi: 10.1016/j.jinf.2024.106251. Epub 2024 Aug 26. PMID: 39182652.
16. Kaal AG, Washer L, Wasylyshyn A, Steyerberg EW, Mahajan P, van Nieuwkoop C. Bacteremia Prediction Models to Reduce Unnecessary Blood Cultures: External Validation in a Large US Emergency Department. [Submitted for publication]
17. Steyerberg EW. (2009) Clinical prediction models: a practical approach to development, validation, and updating. Springer, New York
18. Shapiro NI, Wolfe RE, Wright SB, Moore R, Bates DW. Who needs a blood culture? A prospectively derived and validated prediction rule. *J Emerg Med.* 2008 Oct;35(3):255-64. doi: 10.1016/j.jemermed.2008.04.001. Epub 2008 May 16. PMID: 18486413.
19. Lautenbach E, Localio R, Nachamkin I. Clinicians required very high sensitivity of a bacteremia prediction rule. *J Clin Epidemiol.* 2004 Oct;57(10):1104-6. doi: 10.1016/j.jclinepi.2004.03.003. PMID: 15528062.
20. Dregmans E, Kaal AG, Meziyerh S et al. Analysis of Variation Between Diagnosis at Admission vs Discharge and Clinical Outcomes Among Adults With Possible Bacteremia. *JAMA Netw Open.* 2022 Jun 1;5(6):e2218172. doi: 10.1001/jamanetworkopen.2022.18172. PMID: 35737389; PMCID: PMC9226997.
21. Kaal AG, Op de Hoek L, Hochheimer DT et al. Outcomes of community-acquired pneumonia using the Pneumonia Severity Index versus the CURB-65 in routine practice of emergency departments. *ERJ Open Res.* 2023 May 2;9(3):00051-2023. doi: 10.1183/23120541.00051-2023. PMID: 37143846; PMCID: PMC10152258.
22. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *New England Journal of Medicine.* 1997;336(4):243-250.
23. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax.* 2003;58(5):377-382.
24. van Daalen FV, Boersma WG, van de Garde EMW et al. Management of community-acquired pneumonia in adults: the 2024 practice guideline from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). 2024, SWAB. <https://swab.nl/nl/cap> Accessed on 26th of August, 2025.
25. Kaal A, Snel L, Dane M et al. Diagnostic yield of bacteriological tests and predictors of severe outcome in adult patients with COVID-19 presenting to the emergency department. *Emerg Med J.* 2021 Sep;38(9):685-691. doi: 10.1136/emered-2020-211027. Epub 2021 Jul 21. PMID: 34289966; PMCID: PMC8300553.
26. Kaal AG, Roos R, de Jong P et al. Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study. *Sci Rep.* 2024 Apr 9;14(1):8271. doi: 10.1038/s41598-024-59026-2. PMID: 38594555; PMCID: PMC11004140.
27. Kaal AG, Roos R, Wilms EB, Borgdorff H, van Nieuwkoop C. Geef antibiotica bij voorkeur oraal [Choose oral over intravenous antibiotic therapy]. *Ned Tijdschr Geneesk.* 2024 Oct 15;168:D8300. Dutch. PMID: 39435708.



The background is an abstract composition of warm, vibrant colors including shades of orange, yellow, red, and magenta. The colors are blended together with visible, expressive brushstrokes, creating a sense of movement and depth. The overall effect is a rich, textured surface that serves as a backdrop for the text.

Part I

Diagnostics: Improving the diagnostic process
of bacteremia using procalcitonin



Chapter 2

The diagnostic accuracy of procalcitonin for community-acquired bacteraemia: an updated systematic review and meta-analysis

Anna G. Kaal, Margot Nieberg, Koen Stegmeijer, Ewout W. Steyerberg, Cees van Nieuwkoop

ABSTRACT

Background

Procalcitonin is known to have moderate diagnostic accuracy for bacteremia. A 2014 meta-analysis showed 76% sensitivity for a 0.50 ng/mL threshold. Lower thresholds might improve sensitivity.

Objectives

To determine the diagnostic accuracy of procalcitonin for community-acquired bacteraemia by conducting a systematic review and meta-analysis, focusing on the ability to exclude bacteraemia.

Methods

Data sources: We searched PUBMED, EMBASE and Web of Science from 1 January 2014 to 20 May 2025.

Study eligibility criteria and participants

Articles studying diagnostic accuracy of procalcitonin for community-acquired bacteraemia in adults.

Test and reference standard

Procalcitonin was compared with blood culture results.

Assessment of risk of bias

Risk of bias was assessed using the QUADAS-2 tool.

Methods of data synthesis

We pooled sensitivity/specificity with a bivariate random-effects model and created a summary receiver-operating curve. The main analysis focused on studies reporting on a procalcitonin threshold of 0.10 ng/mL. In addition, we analysed results for all studies, studies with a 0.25 ng/mL and studies with a 0.50 ng/mL threshold.

Results

We included 40 of 5450 identified articles, reflecting 192 529 patients of whom 31 480 (16%) had bacteraemia. Of 40 studies, 32 had high risk of bias. The pooled sensitivity for a 0.10 ng/mL threshold was 93% (95% CI: 85-97%) with a specificity of 36% (95% CI: 26-47%). The area under the summary receiver-operating curve for all studies was 0.80 (95% CI: 0.76-0.83%; prediction interval 0.57-0.91).

Discussion

A low cut-off value of procalcitonin can be useful to exclude community-acquired bacteraemia, depending on what the treating clinician considers to be an acceptable trade-off between sensitivity and specificity. Procalcitonin may require combination with clinical characteristics for accurate assessment of the risk of bacteraemia and safely reducing unnecessary blood cultures.

INTRODUCTION

Community-acquired bacteraemia (CAB) continues to lead to high morbidity and mortality if not treated timely and adequately.¹ Recognizing bacteraemia can be difficult because patients' symptoms are unspecific and most commonly used diagnostic tools either lack accuracy or do not lead to timely results to be used at an emergency department (ED). The fear of missing bacteraemia leads to overuse of diagnostic tests. Currently, the reference standard is blood cultures.² These only yield a true positive result in 4-11% of all blood cultures taken at the ED, whereas another 2-8% is contaminated with skin flora, and performing blood cultures is costly.³⁻⁸ It would therefore be of great value to find a way to improve the diagnostic process for bacteraemia. Because most patients with suspected infection or sepsis at ED do not have bacteraemia, finding a way to immediately rule out bacteraemia would help avoid unnecessary blood cultures and negative consequences of false-positive results.⁶⁻⁸ This is also in line with the increasing call for diagnostic stewardship aiming to minimize inappropriate diagnostic microbiology testing.^{9,10}

Procalcitonin (PCT), a biomarker with growing relevance to identify bacterial infections, has been extensively researched to identify bacteraemia.¹¹⁻¹⁵ A systematic review and meta-analysis from 2007 concluded that PCT had a moderate performance for diagnosing bacteraemia with an area under the curve (AUC) value of 0.84, a pooled sensitivity of 76% (95% CI: 66-84%) and a pooled specificity of 70% (95% CI: 60-79%).¹¹ In 2015, Hoeboer et al.¹² conducted a meta-analysis that concluded that PCT had a fair diagnostic accuracy for bacteraemia with an AUC of 0.78, a pooled sensitivity of 76% (95% CI: 69-82%) and a pooled specificity of 69% (95% CI: 64-72%) in the ED population. Both meta-analyses used cut-offs closest to 0.5 ng/mL in their analysis, whereas especially low values of PCT (<0.10 ng/mL, <0.25 ng/mL) are reported to have high sensitivities.¹²⁻¹⁴

Since this last review, multiple studies have been published that evaluate the diagnostic accuracy of PCT for bacteraemia using lower thresholds. This could lead to a better understanding of the usability of PCT to safely withhold blood cultures. Therefore, we aim to provide an up-to-date overview of the diagnostic accuracy of PCT for bacteraemia in adult patients presenting to the hospital with suspected infection, with a particular focus on its ability to exclude CAB, by performing a systematic review and meta-analysis.

METHODS

Search strategy and study selection

A search strategy was set up in collaboration with an information specialist (PG), see the supplementary material. The search strategy was built on a combination of terms for PCT, bacteraemia and diagnostic properties. We searched PUBMED, EMBASE and Web of Science from 1 January 2014 to 20 May 2025. We used Rayyan to identify potential duplicates. This programme estimates the probability that an article is included in duplicate. We deleted articles with a certainty over 95%. Studies were screened by title and abstract by a minimum of two individual researchers (AGK, MN, KS, CvN), and final inclusion was done after full-text review. Any disagreement was resolved through discussion. Finally, the reference lists of included articles were searched for additional relevant articles.

We included studies that studied solely adult patients ≥ 18 years suspected of having an infection. Articles were eligible for inclusion if they compared the diagnostic accuracy for CAB for PCT compared with the reference standard: blood cultures. They had to report diagnostic properties needed to create a 2×2 table of PCT versus blood culture results. CAB had to be defined as growth of a bacterial pathogen from a blood culture, taken at ED and/or within 48 hours of hospital admission. PCT had to be tested within 24 hours of blood culture withdrawal. Finally, a minimum of one cut-off value for PCT had to be defined. We included possible articles from all languages, using the translation app DeepL for languages other than English, Dutch or Spanish. Exclusion criteria were animal studies, case report or case-control studies, studies studying catheter related bacteraemia, studies that included common skin contaminants in their definition for positive blood culture or studies in burn patients. For studies that seemed eligible but did not report all necessary information, we contacted the corresponding author per e-mail to request additional information. Authors of studies that only reported on part of the cut-offs of interest were also asked to provide additional results. This study is reported according to the preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA) standards.¹⁶

Study protocol

We registered the study protocol at Prospero, with registration number CRD42023426791.

Data collection

We used a predefined data collection form (see supplementary material) for data collection. All information was collected in duplicate. Any disagreement was resolved by discussion, and in case of persistent disagreement, one of the other authors was consulted. We collected data regarding study characteristics (study design, country, and study year), specific information regarding the patient population such as reported use of immunosuppressants or recent antibiotics, whether there was a clear definition of false-positive blood cultures and the method used to analyse PCT.

Quality assessment

We used the QUADAS-2 tool to evaluate risk of bias and concerns regarding applicability.¹⁷ All evaluations were done in duplicate, with discrepancies resolved by discussion.

Statistical analysis

Because we aim to analyse the ability of low PCT values to exclude bacteraemia, our main analysis focused on the results reported for a PCT threshold of 0.10 ng/mL, including lower thresholds and up to 0.15 ng/mL. In addition, we analysed results for (1) thresholds of 0.25 ng/mL (including 0.15 ng/mL up to 0.35 ng/mL), (2) the more commonly suggested threshold of 0.50 ng/mL, including thresholds from 0.35 up to 0.65 ng/mL and (3) using the results of one PCT cut-off closest to 0.5 ng/mL for all included studies. For each analysis, only one cut-off per study was included.

We used the bivariate random-effects regression model to pool sensitivity and specificity. This was also used to create a summary receiver-operating curve for all included studies, using the results of the reported threshold closest to 0.5 ng/mL for comparability with previous studies. We included a prediction region to reflect the estimated sensitivity and specificity for a new study, taking between-study heterogeneity into account. We do not report quantitative measures of heterogeneity because these are not routinely used in DTA reviews because they do not account for heterogeneity explained by phenomena such as different positivity thresholds between studies.¹⁸ Heterogeneity can be visually inspected by the degree to which the observed study results lie close to the summary ROC curve. We estimated the prediction interval for the AUC of the ROC curve by running a random-effect meta-analysis with the 'metafor' package on all studies that reported the AUC with CI, estimating the results on a logit scale.¹⁹ In addition, we explored the influence of study and patient characteristics by performing multiple subanalyses in selected subgroups: prospective versus retrospective studies, according to setting (ED, ward, intensive care unit, or mix), studies with a clear definition of false-positive blood

cultures versus those without, studies including immunocompromised patients versus studies with immunocompetent patients only, according to whether studies included patients with recent use of antibiotics and, finally, according to specific site of infection. We used Deeks' funnel plot asymmetry test to evaluate publication bias, using the commonly used p value of <0.10 for the slope coefficient as a sign of significant asymmetry, indicating potential publication bias.¹⁸ A sensitivity analysis was done by calculating the summary estimates for low and high risk of bias studies and low level of concern and high level of concern for applicability as assessed by the QUADAS-2 criteria. Unclear risk was reclassified as high risk for this comparison. All analyses were conducted using Rstudio version 2024.09.0.

RESULTS

Study selection

Our literature search retrieved 5450 unique articles, of which 5255 were excluded after screening of title and abstract. A total of 191 full-text articles were screened for eligibility, of which 151 were excluded. Reasons for exclusion are depicted in Fig. 1; 31 were excluded because of multiple reasons that are separately listed in the supplementary material. Finally, a total of 40 cohort studies were included (Fig. 1, Table 1) accounting 192 529 patients of whom 31 480 (16.3%) had bacteraemia. The reported PCT cut-offs ranged from 0.10 to 10.83 ng/mL. All studies used automatized PCT assays (see supplementary material).

Risk of bias and applicability

We evaluated all studies using the QUADAS-2 tool (Table 2). Overall, 32 of 40 studies had a high risk of bias on at least one domain. An important source of bias was the unclear handling of contaminated blood cultures, by either not using a clear definition or unclear handling of the contaminated cultures in the analysis. We found less applicability concerns, only in 9 of 40 studies, which were mostly caused by the selection of a very specific population.

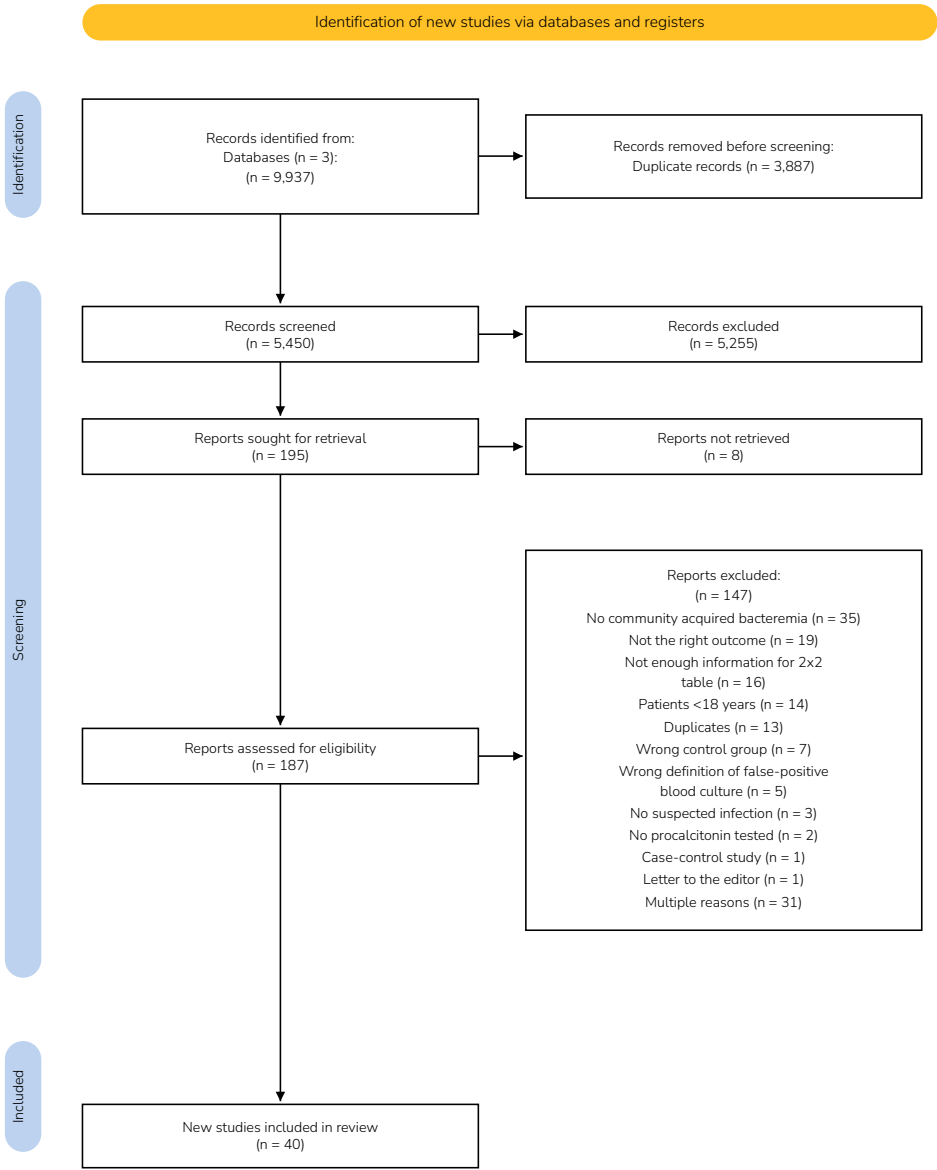


Figure 1. Flowchart of identification of new studies

Table 1. Characteristics of all included studies.

Study	Country	Type	Inclusion criteria	Setting	Patients (n)	Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised	Antibiotic treatment	Cut-off values (ng/mL)
Aziji et al., 2022 ²⁰	Netherlands	Pro-spective	Suspected infection with performing of blood cultures and influenza PCR test	ED	546	64 (17)	301 (55)	47 (9)	Clear	Viral	Yes	Yes	0.1, 0.25, 0.5
Choe et al., 2016 ²¹	Korea	Retro-spective	Septic shock or severe sepsis with blood lactate ≥2.0 mmol/L	ED	1212	66 (55-74)	718 (59)	421 (35)	Unclear	Mixed	Yes	Unclear	0.25, 0.50
Ciotti et al., 2019 ²²	Italy	Pro-spective	Febrile outpatients with haematological malignancies	ED	70	Not specified	Not specified	11 (16)	Unclear	Mixed	Yes	Unclear	0.5
García de Guadiana-Romualdo et al., 2019 ²³	Spain	Pro-spective	Febrile neutropenia	ED	111	63 (21-85)	42 (38)	16 (14)	Clear	Mixed	Yes	Unclear	0.34
Julián-liménez et al., 2022 ²⁴	Spain	Pro-spective	Patients >65 y with clinically suspected infection	ED	2401	80 (8)	1489 (62)	579 (24)	Clear	Mixed	Unclear	Unclear	0.5
Harikrishna et al., 2020 ²⁵	India	Pro-spective	Acute fever	Mixed	450	51 (18)	312 (65)	44 (10)	Clear	Mixed	Unclear	Unclear	0.5
Hertz et al., 2024 ²⁶	Denmark	Pro-spective	Suspected urogenital tract infection	ED	173	75 (17)	102 (59)	39 (23)	Clear	Urogenital	Unclear	Unclear	0.15
Hoenigl et al., 2014 ²⁷	Austria	Pro-spective	Positive SIRS criteria	Mixed	515	BC-: 68, BC+: 66a	519 (58)a	282 (55)	Clear	Mixed	Unclear	No	0.1, 0.25, 0.5

Study	Country	Type	Inclusion criteria	Setting	Patients (n)	Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised	Antibiotic treatment	Cut-off values (ng/mL)
Inoue et al., 2020 ²⁸	Japan	Unclear	Diagnoses of infection and hospitalisation after performing blood cultures	ED	5064	77 (68-83)	3217 (64)	365 (7)	Unclear	Mixed	Unclear	Unclear	0.45
Iqbal et al., 2020 ²⁹	Spain	Retro-spective	Suspected infection	ED	1425	53 (19)	666 (47)	179 (13)	Clear	Mixed	Yes	Yes	0.51
Jimenez-Santos et al., 2014 ³⁰	Spain	Pro-spective	Oncology patients with fever	ED	152	63 (22)	71 (53)	22 (14)	Clear	Mixed	Yes	Unclear	0.11, 0.2, 0.5
Julían-Jiménez et al., 2022 ³¹	Spain	Pro-spective	Clinical suspicion and diagnosis of infectious processes	ED	4436	67 (19)	2648 (60)	899 (20)	Clear	Mixed	Yes	Yes	0.1, 0.25, 0.51
Kaal et al., 2024 ³²	Netherlands	Pro-spective	Performing blood cultures	ED	2111	63 (NA)	1147 (54)	273 (13)	Clear	Mixed	Yes	Yes	0.1, 0.25, 0.5
Kim et al., 2015 ³³	Korea	Retro-spective	Testing of ANC, CRP, PCT and body temperature	ED	3305	Not specified	1855 (56)	564 (17)	Clear	Mixed	Unclear	Unclear	0.1
Laukemann et al., 2015 ¹⁴	Switzerland	Pro-spective	Suspected infection	ED	1083	67 (53-78)	623 (58)	104 (10)	Clear	Mixed	Yes	Yes	0.1, 0.25, 0.5
Lawandi et al., 2023 ³⁴	USA	Retro-spective	Blood cultures and PCT testing within 24 h of admission	Mixed	74958	Not specifiedb	36589 (50)	9865 (13)	Clear	Mixed	Unclear	Unclear	0.1, 0.25, 0.5
Lee et al., 2018 ³⁵	Zuid Korea	Retro-spective	Suspected acute cholangitis	ED	204	76 (17-102)	118 (58)	65 (32)	Unclear	Abdominal	Unclear	Unclear	0.5

Study	Country	Type	Inclusion criteria	Setting (n)	Patients Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised	Antibiotic treatment	Cut-off values (ng/mL)	
Leroux et al., 2021 ³⁶	France	Retro-spective	PCT testing	ED	852	62 (23)	419 (49)	74 (9)	Clear	Mixed	Unclear	0.1, 0.25, 0.5	
Lin et al., 2017 ³⁷	Taiwan	Retro-spective	Performing blood cultures	ED	886	Not specified	197 (22)	Clear	Mixed	Unclear	Unclear	0.5	
Ljungstrom et al., 2017 ³⁸	Sweden	Pro-spective	Suspected community-onset sepsis	ED	1572	71 (58-81)	875 (56)	197 (13)	Unclear	Mixed	Unclear	0.1	
Loonen et al., 2014 ³⁹	Netherlands	Retro-spective	Clinically suspected infection with >2 SIRS criteria and performing of blood cultures	ED	118	BC-: 60, BC+: 69	74 (59)	20 (17)	Clear	Mixed	Unclear	2	
Marik et al., 2020 ⁴⁰	USA	Retro-spective	Testing of PCT, lactate, CBC with differential and blood cultures for suspected bacteraemia	Mixed	1767	61 (49-71)	NA (52)	384 (22)	Clear	Mixed	Unclear	0.5	
Naffaa 2014 ⁴¹	Israel	Pro-spective	Sepsis of any cause	Mixed	40	79 (68-84)	25 (63)	10 (25)	Clear	Mixed	No	1.35	
Nishino 2017 ⁴²	Japan	Pro-spective	Acute cholangitis	Mixed	62	75	46 (74)	29 (47)	Unclear	Abdominal	Unclear	3	
Oksuz et al., 2015 ⁴³	Turkey	Pro-spective	Febrile patients >37°C	Mixed	661	52 (18)	Not specified	77 (12)	Clear	Mixed	Unclear	0.1, 0.25, 0.5	
Oura et al., 2023 ⁴⁴	Japan	Retro-spective	Suspected acute cholangitis	Mixed	127	77 (71-85)	93 (67)c	71 (56)	Clear	Abdominal	Unclear	Yes	2.8

Study	Country	Type	Inclusion criteria	Setting	Patients (n)	Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised treatment	Antibiotic Cut-off values (ng/mL)
Palma et al., 2024 ⁴⁵	Spain	Retro-spective	Patients ≥75 y with clinically suspected infection and performing of blood cultures	ED	109	83 (6)	64 (59)	22 (20)	Clear	Mixed	Yes	Unclear 0.76
Romualdo et al., 2014 ⁴⁶	Spain	Pro-spective	At least two SIRS criteria and clinical suspicion of bacteraemia with performing blood cultures	ED	226	67 (26)	132 (58)	37 (16)	Clear	Mixed	Unclear	Yes 0.45
Roy et al., 2018 ⁴⁷	USA	Retro-spective	Suspected bacterial infection, PCT and blood cultures tested within 24 h of admission	Mixed	339	66 (17)	33 (50)	25 (7)	Clear	Mixed	Unclear	Unclear 0.5
Seo et al., 2015 ⁴⁸	Japan	Pro-spective	Admission due to infection	Mixed	78	79	40 (51)	29 (37)	Unclear	Mixed	No	Unclear 0.1, 0.25, 0.5
Takahashi et al., 2024 ⁴⁹	Japan	Retro-spective	Admission due to DKA or HHS	Mixed	608	58 (19)	329 (54)	70 (12)	Clear	Mixed	Unclear	Yes 0.1, 0.25, 0.5
Timbrook et al., 2023 ⁵⁰	USA	Retro-spective	Performing of blood cultures and PCT on day one of admission	Mixed	71105	68 (56-79)	35713 (50)	14846 (21)	Clear	Mixed	Yes	Unclear 0.5

Study	Country	Type	Inclusion criteria	Setting (n)	Patients Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised	Antibiotic treatment	Cut-off values (ng/mL)
Toreh et al., 2017 ⁵¹	Indonesia	Unclear	Suspected urosepsis with performing of blood cultures	Mixed 141	Not specified	Not specified	69 (49)	Clear	Urogenital	Unclear	Unclear	0.5
Uyar et al., 2023 ⁵²	Turkey	Pro-spective	Suspected sepsis	ED 47	BC-: 61. BC+: 62	31 (66)	20 (43)	Clear	Mixed	No	Yes	0.55
Wiwatharagoses and Kingnakom 2016 ⁵³	Thailand	Pro-spective	Symptoms of systemic infections such as sepsis, severe sepsis, or septic shock	ED 110	52 (21)	59 (54)	23 (21)	Clear	Mixed	No	No	2
Xiao et al., 2024 ⁵⁴	China	Retro-spective	Admission due to sepsis or septic shock with testing blood cultures, PCT, presepsin and CRP	Mixed 577	70 (59-81)	340 (59)	170 (29)	Unclear	Mixed	Yes	Unclear	4.31
Yamaguchi et al., 2023 ⁵⁵	Japan	Retro-spective	Suspected acute cholangitis with performing of blood cultures	Mixed 69	BC-: 75. BC+: 82	44 (47)d	35 (51)	Clear	Abdominal	No	No	0.5
Youkhana et al., 2017 ⁵⁶	USA	Retro-spective	PCT and blood cultures performed within 24 h	Mixed 14387	Not specified	Not specified	1160 (8)	Clear	Mixed	Yes	Unclear	0.5

Study	Country	Type	Inclusion criteria	Setting	Patients (n)	Age (y)	Male (%)	True positive BCs (%)	Contamination	Infection	Immuno-compromised treatment	Antibiotic values (ng/mL)	Cut-off values (ng/mL)
Zhang et al., 2021 ⁵⁷	China	Pro-spective	At least two SIRS criteria with clinical suspicion of bacteraemia and performing blood cultures	Mixed	149	BC-: 59, BC+: 57	93 (62)	59 (40)	Clear	Mixed	No	Unclear	1.29
Zhang et al., 2023 ⁵⁸	China	Pro-spective	Suspected acute cholangitis and performing of prepspsin, acylcarnitines and blood cultures	Mixed	280	74 (66-84)	166 (59)	92 (33)	Unclear	Abdominal	Unclear	Unclear	10.83

Type = retrospective versus prospective cohort study. Contamination: was there a clear definition of contaminated blood cultures? Immunocompromised: were immunocompromised patients included? Antibiotic treatment: inclusion of patients with recent (<7 d) antibiotic treatment. ANC, absolute neutrophil count; BC, blood culture; CBC, complete blood count; CRP, C-reactive protein; DKA, diabetic keto-acidosis; ED, emergency department; HHS, hyperglycemic hyperosmolar state; PCT, procalcitonin; SIRS, systemic inflammatory response syndrome.

a A total of 515 patients had a BC and PCT test; the mean age and gender were only given for the total number of patients: 898.

b Only age given for six subgroups, with medians ranging between 65 and 68 y.

c A total of 127 patients had a BC and PCT test; the mean age was only given for the total number of patients: 138.

d A total of 69 patients had a BC and PCT test; the mean age was only given for the total number of patients: 94.

Table 2. Risk of bias and applicability concerns per individual study as assessed by the QUADAS-2 tool

	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Azjili et al., 2022 ²⁰	Low	Low	Low	Low	Low	Low	Low
Choe et al., 2016 ²¹	High	Low	High	High	Low	Low	High
Crotti et al., 2019 ²²	Unclear	Low	Unclear	Unclear	High	Low	Low
García de Guadiana-Romualdo et al., 2019 ²³	Low	High	Low	Low	Low	Low	Low
Julían-Jiménez et al., 2022 ²⁴	High	Low	Low	Low	Low	Low	Low
Harikrishna et al., 2020 ²⁵	Low	High	High	Low	Low	Low	Low
Hertz et al., 2024 ²⁶	Low	High	Low	High	Low	Low	Low
Hoenigl et al., 2014 ²⁷	High	Low	Low	Low	Low	Low	Low
Inoue et al., 2020 ²⁸	Unclear	High	Unclear	Unclear	Unclear	Unclear	Unclear
Iqbal et al., 2020 ²⁹	High	Low	Low	Low	Low	Low	Low
Jimenez-Santos et al., 2014 ³⁰	Low	Unclear	Low	Low	Low	Low	Low
Julían-Jiménez et al., 2022 ³¹	Low	Low	Low	High	Low	Low	Low
Kaal et al., 2024 ³²	Low	Low	Low	Low	Low	Low	Low
Kim et al., 2015 ³³	Low	Low	Low	Low	Low	Low	Low
Laukemann et al., 2015 ¹⁴	Low	Low	Low	Low	Low	Low	Low
Lawandi et al., 2023 ³⁴	High	Low	Low	High	Low	Low	Low
Lee et al., 2018 ³⁵	Low	High	High	Low	Low	Low	Low
Leroux et al., 2021 ³⁶	Low	Low	High	High	Low	Low	Low
Lin et al., 2017 ³⁷	Low	Low	Low	Low	Low	Low	Low
Ljungstrom et al., 2017 ³⁸	Low	Low	High	High	Low	Low	Low
Loonen et al., 2014 ³⁹	Low	Unclear	Low	High	Low	Low	Low

	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Marik et al., 2020 ⁴⁰	Low	Low	Low	High	Low	Low	Low
Naffaa 2014 ⁴¹	Low	Low	High	Unclear	Low	Low	Low
Nishino 2017 ⁴²	Unclear	High	Unclear	Unclear	Low	Low	Unclear
Oksuz et al., 2015 ⁴³	Low	High	Low	Low	Low	Low	Unclear
Oura et al., 2023 ⁴⁴	Low	Low	Low	Low	Low	Low	Low
Palma et al., 2024 ⁴⁵	Low	High	Low	Low	Low	Low	Low
Romualdo et al., 2014 ⁴⁶	Low	High	Low	Low	Low	Low	Low
Roy et al., 2018 ⁴⁷	Unclear	Low	High	Unclear	Low	Low	Low
Seo et al., 2015 ⁴⁸	High	High	High	High	High	Low	High
Takahashi et al., 2024 ⁴⁹	Unclear	Low	Low	High	High	Low	Low
Timbrook et al., 2023 ⁵⁰	Low	Low	Low	High	Low	Low	Low
Toreh et al., 2017 ⁵¹	Unclear	Low	Unclear	Unclear	Low	Low	Unclear
Uyar et al., 2023 ⁵²	Low	High	Low	High	Low	Low	Low
Wiwatcharagoses and Kingnakom 2016 ⁵³	High	Low	High	High	Low	Low	Low
Xiao et al., 2024 ⁵⁴	High	High	High	Unclear	Low	Low	Low
Yamaguchi et al., 2023 ⁵⁵	High	Low	Low	High	Low	Low	Low
Youkhana et al., 2017 ⁵⁶	Unclear	Unclear	Low	High	Unclear	Unclear	Unclear
Zhang et al., 2023 ⁵⁸	High	High	High	Unclear	Low	Low	Low
Zhang et al., 2021 ⁵⁷	Low	High	Low	High	Low	Low	Low

Results of individual studies

Of all studies reporting on the lowest threshold (up to 0.15 ng/mL), sensitivity ranged from 28% to 99.9% with corresponding specificity of 15-80%. Fig. 2 shows the forest plot of all included studies. The corresponding 2 × 2 table can be found in the supplementary material. Of 40 studies, 30 reported an AUC of the ROC curve, ranging from 0.56 to 0.96, with substantial uncertainty per study.

Pooled results

The pooled sensitivity for the 0.10 ng/mL threshold was 93% (95% CI: 85-97%) with a specificity of 36% (95% CI: 26-47%) and an AUC of 0.68 (95% CI: 0.53-0.80). At the average bacteraemia prevalence of 16.3%, this would lead to a negative predictive value of 97.1%. The sensitivity prediction interval for a 0.10 ng/mL cut-off was 24-99.8% with specificity ranging between 17% and 94%. See Table 3 for the results for the other thresholds. The pooled AUC for all studies was 0.80 (95% CI: 0.76-0.83) (Fig. 3). The prediction interval for the AUC was 0.57-0.91.

Table 3. Pooled sensitivity and specificity with corresponding prediction intervals for all four analyses.

PCT threshold	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Sensitivity prediction interval (%)	Specificity prediction interval (%)
0.10 ng/mL	92.8 (85.3-96.7)	35.7 (26.1-46.7)	24.0-99.8	17.0-94.0
0.25 ng/mL	85.1 (74.0-92.0)	55.4 (43.4-66.9)	20.0-99.0	8.0-88.0
0.50 ng/mL	78.6 (71.8-84.2)	68.9 (59.3-77.1)	27.0-97.0	3.0-87.0
Closest to 0.50 ng/mL	78.2 (72.8-82.7)	69.1 (62.0-75.7)	30.0-97.0	4.0-84.0

The first three analyses include only articles reporting on that threshold, whereas the fourth includes all articles using the threshold closest to a procalcitonin value of 0.50 ng/mL. The pooled sensitivity and specificity reflect the estimated average result. The prediction interval answers the question: 'If I would perform a new study, what will be my sensitivity and specificity with 95% certainty?

Subanalyses

In the subanalyses of all groups, pooled sensitivity was higher for an ED setting than a mixed setting (95% vs. 89% at a 0.10 ng/mL cut-off, prediction interval 42-99.8%). Studies using a clear definition of contaminated blood cultures also showed higher sensitivity than those without (80% vs. 69%, cut-off closest to 0.50 ng/mL). In the sensitivity analysis, we found higher pooled sensitivity for studies with low risk of bias, especially for the lower cut-offs (97% vs. 87% at a 0.10 ng/mL cut-off, specificity 34% vs. 37%, negative predictive value 98.3% vs. 93.6%, supplementary material). The prediction intervals for these low bias studies were 58-99.9% for sensitivity and 28-90% for specificity.

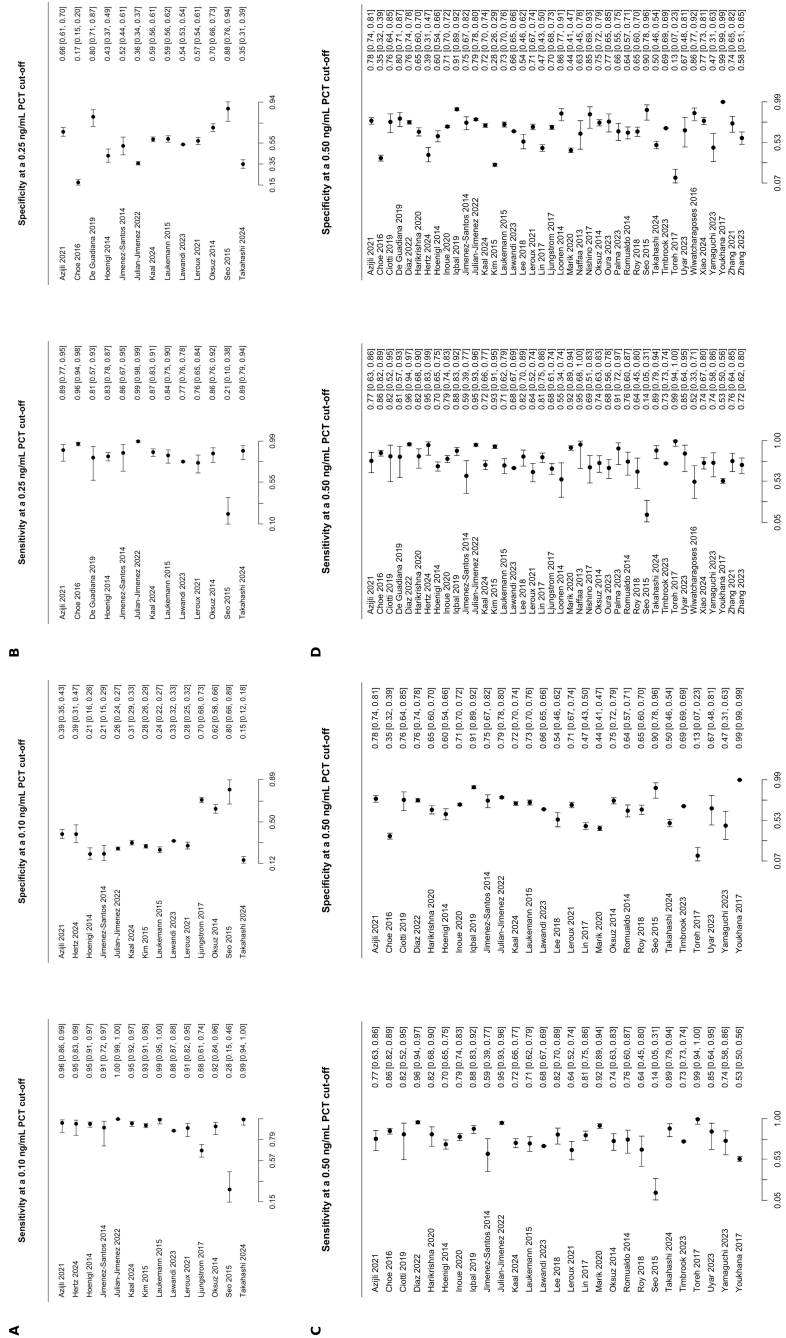


Fig. 2: Coupled forest plots of sensitivity and specificity of procalcitonin for the prediction of bacteraemia for all included studies. The four different panels A-D represent different cut-offs of procalcitonin. Panel (A) displays results for a cut-off of 0.10 ng/mL, (B) for 0.25 ng/mL, (C) for 0.50 ng/mL and (D) included all studies with the cut-off closest to 0.50 ng/mL.

Publication bias

There was no indication of publication bias based on Deeks' funnel plot (p 0.59) (see supplementary material).

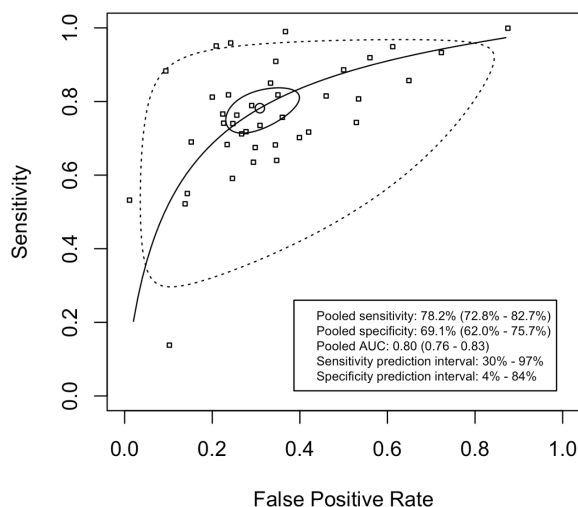


Fig. 3: Summary receiver operating characteristic (sROC) curve for the diagnostic accuracy of procalcitonin for bacteraemia, including the results at the reported threshold closest to 0.50 ng/mL. The small circle reflects the pooled sensitivity and specificity with the 95% CI shown as the larger solid circle, whereas the squares represent the results of the individual studies. The dotted line represents the prediction region, which illustrates the expected variability of sensitivity and specificity of a new study, accounting for between-study heterogeneity.

DISCUSSION

In this study, we updated a systematic review and meta-analysis of all studies published since the last meta-analysis regarding the diagnostic accuracy of PCT for CAB. We chose a new approach, focusing on the potential use of PCT to exclude CAB, to safely withhold unnecessary blood cultures. We found that with a low cut-off of 0.10 ng/mL, for a typical average study, the rate of missed bacteraemic episodes would be 7.2% (95% CI: 3.3-15%), while saving 36% (95% CI: 26-47%) of blood cultures. At the ED, the expected potential average savings are comparable (34% [95% CI: 25-44%]) with around 5% missed true positive blood cultures (95% CI: 2-11%).

Our findings are in line compared with one previous meta-analysis regarding a 0.50 ng/mL cut-off: a pooled sensitivity of 79% versus 76% and pooled specificity of

69% versus 69%, and an AUC of 0.81 versus 0.79¹². Another meta-analysis also showed a comparable AUC of 0.84 (0.75-0.90).¹¹

It is important to recognize the difference between the estimated pooled parameters with their CIs and the prediction intervals. The first estimates an average value. However, the prediction interval answers the question: 'If I would perform a new study, what will be my sensitivity and specificity with 95% certainty?'. This might be more relevant for clinical practice, and our results show that there is substantial between-study variation. The sensitivity prediction interval for a 0.10 ng/mL cut-off was very large (24-99.8%), as was the prediction interval for the AUC, ranging from 0.57 to 0.91. However, it is known that large heterogeneity leads to irreducible uncertainty, even when more studies would be performed.¹⁹ We did observe smaller intervals for studies with low risk of bias and ED studies. This underlines that the usability of PCT for excluding bacteraemia is highly dependent on the intended clinical setting.

On the basis of our findings, we believe that PCT can indeed be useful to rule out CAB in a typical setting, but we note that its value will largely depend on what risk of missing a true positive blood culture the treating clinician would find acceptable. It is known that clinicians require a high sensitivity, because this implies that we can safely exclude bacteraemia.⁵⁹ However, by decreasing the PCT threshold to increase sensitivity, specificity will inevitably decrease. This leads to less reduction of blood cultures, and will lead to more false-positive cultures. Some clinicians might argue that missing around 7% is unacceptable. The question remains whether this caution is justified because the vast majority of blood cultures do not impact treatment decisions; only 0.2-3% of all blood cultures taken at the ED do.^{60,61} Meanwhile, it is known that false-positive blood cultures are common (2-8%) and associated with negative patient outcomes such as increased length of stay, increased costs and more exposure to antibiotics.³⁻⁸ Therefore, we would argue that the local rate of false-positive cultures might be a good starting point when thinking about what rate of missed true positive blood cultures is acceptable.

In this meta-analysis, we choose to create four groups reflecting four thresholds. Each analysis was done using the bivariate model, including one threshold result per study. The Cochrane Handbook for diagnostic test accuracy reviews recommends using the hierarchical summary receiver operating characteristic (HSROC) model when studying multiple thresholds, because this model can include multiple thresholds per study in one analysis.¹⁸ On the basis of this model, one could estimate the pooled sensitivity and specificity for a 0.10 ng/mL cut-off. However, given our primary interest in the results of this lower threshold, we elected to analyse studies that reported results for this threshold, as opposed to estimating results primarily

based on data from higher thresholds. In addition, we have opted for a similar approach to that employed in the 2014 meta-analysis. This might help to increase the comparability of results.¹²

PCT is not only used as a diagnostic test, but is also extensively studied as a biomarker to guide antibiotic treatment. A meta-analysis showed that PCT-based algorithms reduced antibiotic exposure without an increase in mortality.⁶² Interestingly, although most of these PCT algorithms advice to withhold or to discontinue antibiotic treatment at a PCT level below 0.25 ng/mL, we showed that around 15% of true positive blood cultures did not have a PCT value over 0.25 ng/mL at presentation. This is a finding clinicians should be aware of when considering implementing such an algorithm.

It is impressive how many new articles in the past 10 years studied the diagnostic accuracy of PCT for CAB, whereas the results are comparable with previous meta-analysis. We believe that further studies regarding diagnostic accuracy of PCT for CAB will not lead to substantial reduction of uncertainty. Future studies should focus on either reporting even lower cut-offs of PCT that might lead to an acceptable trade-off between sensitivity and specificity, or on the interpretation of PCT in the clinical context. Prediction models that combine PCT with clinical predictors such as age and vital signs might help to safely withhold blood cultures.³² Finally, studies focusing on the cost-effectiveness of PCT testing for CAB will also help to determine its use in clinical practice.⁶³

We did not find an indication for publication bias using Deeks' funnel plot. However, detecting publication bias for diagnostic test accuracy studies is notoriously difficult, and cannot be excluded based on this test.¹⁸ However, we performed a very extensive search across multiple databases to retrieve potential studies, including studies in all languages, and reached out to authors in cases of missing data or unclear data. These efforts have likely decreased the risk of publication bias in this meta-analysis.

Notably, we found that studies using a clear definition of contaminated blood cultures found higher sensitivities. This makes sense, because false-positive blood cultures do not represent an inflammatory process and PCT values will likely not be elevated. If these cultures are analysed as true positives, this would falsely lead to a lower sensitivity. We would argue that when conducting research regarding blood cultures, defining a false-positive blood culture and analysing them as negative cultures is essential to produce reliable results. In this analysis, 15 of 40 included articles failed to do so, leading to possible biased results.

A strength of this study is the rigorous literature search, in which we independently screened 5450 articles in duplicate. We were able to include 40 articles representing a very large number of patients from a wide range of counties, making the results very robust and broadly applicable. A limitation of the study is that we choose to include the studies with high risk of bias in our main analysis. Studies with a low risk of bias had a higher pooled sensitivity of 97% (95% CI: 92-99%) with a pooled specificity of 34% (95% CI: 26-44%) (please see supplementary material). Our general conclusion might therefore be an underestimation of the true diagnostic accuracy of PCT for CAB. Furthermore, the included studies used a variety of PCT assays. Although all of the assays had a detection range that should be adequate for this study, it cannot be excluded that differences in sensitivity of the used assays will have influenced our results. However, we do not expect this to have led to clinically relevant differences. Finally, we choose the bivariate model in a selection of studies rather than using the hierarchical summary receiver operating characteristic model. This approach leads to a loss of available information in the analysis. However, as explained previously, we consider our approach more appropriate to answer our primary research question focusing on ruling out CAB.

In conclusion, using lower cut-off values of PCT to exclude CAB might be useful, depending on what the treating clinician considers to be an acceptable trade-off between sensitivity and specificity. Studies reporting on the combination of PCT with clinical parameters are needed to increase our understanding of the possible use of PCT to safely withhold unnecessary blood cultures.

CRedit authorship contribution statement

Anna G. Kaal: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. **Margot Nieberg:** Investigation, Validation, Writing—review and editing. **Koen Stegmeijer:** Investigation, Validation, Writing—review and editing. **Ewout W. Steyerberg:** Conceptualization, Methodology, Supervision, Writing—review and editing. **Cees van Nieuwkoop:** Conceptualization, Investigation, Methodology, Supervision, Validation, Writing—review and editing. N/A: Funding acquisition, Resources.

Transparency declaration

Potential conflict of interest

The authors declare that they have no conflicts of interest.

Financial report

No external funding was used to support this study.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors made use of ChatGPT for the sole purpose of enhancing its readability and linguistic quality. After using this tool, the authors have reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgements

We would like to thank Pema Gurung for her help with setting up the search strategy.

SUPPLEMENTARY MATERIAL



REFERENCES

1. Jiménez-Aguilar P, López-Cortés LE, Rodríguez-Baño J. Impact of infectious diseases consultation on the outcome of patients with bacteraemia. *Ther Adv Infect Dis* 2019;6:2049936119893576.
2. Rhodes A, Evans LE, Alhazzani W, et al. Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med* 2017;43:304-77.
3. Long B, Koyfman A. Best clinical practice: Blood culture utility in the emergency department. *J Emerg Med* 2016;51:529-39.
4. Coburn B, Morris AM, Tomlinson G, Detsky AS. Does this adult patient with suspected bacteremia require blood cultures? *JAMA* 2012;308:502-11.
5. Nannan Panday R. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. *PLoS One* 2019;14:e0214052.
6. Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: Importance, implications, interpretation and prevention. *Clin Microbiol Infect* 2018;24:964-9.
7. Dempsey C, Skoglund E, Muldrew KL, Garey KW. Economic health care costs of blood culture contamination: A systematic review. *Am J Infect Control* 2019;47:963-7.
8. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization. The true consequences of false-positive results. *JAMA* 1991;265:365-9.
9. Curren EJ, et al. Advancing diagnostic stewardship for healthcare associated Infections, antibiotic resistance, and sepsis. *Clin Infect Dis* 2022;74:723-8.
10. Morgan DJ, Malani PN, Diekema DJ. Diagnostic stewardship to prevent diagnostic error. *JAMA* 2023;329:1255-6.
11. Jones AE, Fiechtl JF, Brown MD, Ballew JJ, Kline JA. Procalcitonin test in the diagnosis of bacteremia: A meta-analysis. *Ann Emerg Med* 2007;50:34-41.
12. Hoeboer S, van der Geest PJ, Nieboer D, Groeneveld ABJ. The diagnostic accuracy of procalcitonin for bacteraemia: A systematic review and meta-analysis. *Clin Microbiol Infect* 2015;21:474-81.
13. van Nieuwkoop C, et al. Procalcitonin reflects bacteremia and bacterial load in urosepsis syndrome: A prospective observational study. *Crit Care* 2010;14:R206.
14. Laukemann S, et al. Can we reduce negative blood cultures with clinical scores and blood markers? Results from an observational cohort study. *Medicine (Baltimore)* 2015;94:e2264.
15. Goodlet KJ, Cameron EA, Nailor MD. Low sensitivity of procalcitonin for bacteremia at an academic medical center: A cautionary tale for antimicrobial stewardship. *Open Forum Infect Dis* 2020;7:ofaa096.
16. Salameh J, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): Explanation, elaboration, and checklist *BMJ*. 2020;370:m2632.
17. Whiting PF, et al. QUADAS-2 Group. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36.
18. Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. The Cochrane Collaboration; 2023.

19. van Leeuwen FD, et al. Instability of the AUROC of clinical prediction models. *Stat Med* 2025;44:e70011.
20. Azijli K, et al. Optimal use of procalcitonin to rule out bacteremia in patients with possible viral infections. *J Am Coll Emerg Physicians Open* 2022;3(3):e12621.
21. Choe EA, et al. The prevalence and clinical significance of low procalcitonin levels among patients with severe sepsis or septic shock in the emergency department. *Shock* 2016;46:37-43.
22. Ciotti G, et al. Procalcitonin in the diagnostic work-up of febrile episodes in outpatients with hematologic malignancies. *Haematologica* 2019;104:150.
23. García de Guadiana-Romualdo L, et al. Prognostic value of procalcitonin and lipopolysaccharide binding protein in cancer patients with chemotherapy-associated febrile neutropenia presenting to an emergency department. *Biochem Med (Zagreb)* 2019;29:010702.
24. Julián-Jiménez A, et al. Utilidad del modelo 5MPB-Toledo para predecir bacteriemia en el paciente anciano. *Infectio* 2022;26:128-36.
25. Harikrishna J, et al. Serum procalcitonin as a biomarker of bloodstream infection & focal bacterial infection in febrile patients. *Indian J Med Res* 2020;151:342-9.
26. Hertz MA, et al. The diagnostic accuracy of procalcitonin, soluble urokinase-type plasminogen activator receptors, and C-reactive protein in diagnosing urinary tract infections in the emergency department. *J Clin Med* 2024;13:1776.
27. Hoenigl M, et al. Procalcitonin fails to predict bacteremia in SIRS patients: A cohort study. *Int J Clin Pract* 2014;68:1278-81.
28. Inoue H, et al. Procalcitonin (PCT) levels aids in distinguishing extended-spectrum beta-lactamases (ESBLs) producing Enterobacteriaceae bacteremia in emergency department. *J Infect Public Health* 2020;13:317-73.
29. Iqbal-Mirza SZ, et al. Factores predictores de bacteriemia en los pacientes atendidos en el Servicio de Urgencias por infección. *Rev Esp Quimioter* 2020;33:32-43.
30. Jimenez-Santos E, et al. Utilidad de la procalcitonina para predecir bacteriemia en pacientes oncológicos con fiebre. *Rev Lab Clín* 2014;7:127-33.
31. Julián-Jiménez A, et al. A bacteraemia risk prediction model: development and validation in an emergency medicine population. *Infection* 2022;50:203-21.
32. Kaal AG, et al. Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model. *J Infect* 2024;89:106251.
33. Kim SY, et al. Procalcitonin in the assessment of bacteraemia in emergency department patients: Results of a large retrospective study. *Ann Clin Biochem* 2015;52(Pt 6):654-9.
34. Lawandi A, et al. Reliability of admission procalcitonin testing for capturing bacteremia across the sepsis spectrum. *Crit Care Med* 2023;51:1527-37.
35. Lee YS, et al. Procalcitonin as a decision-supporting marker of urgent biliary decompression in acute cholangitis. *Dig Dis Sci* 2018;63:2474-9.
36. Leroux P, et al. Clinical outcome predictive value of procalcitonin in patients suspected with infection in the emergency department. *Emerg Med Int* 2021;2021:2344212.

37. Lin CT, et al. Diagnostic value of serum procalcitonin, lactate, and high-sensitivity C-reactive protein for predicting bacteremia in adult patients in the emergency department. *PeerJ* 2017;5:e4094.
38. Ljungström L, et al. Diagnostic accuracy of procalcitonin, neutrophil-lymphocyte count ratio, C-reactive protein, and lactate in patients with suspected bacterial sepsis. *PLoS One* 2017;12:e0181704.
39. Loonen AJ, et al. Biomarkers and molecular analysis to improve bloodstream infection diagnostics in an emergency care unit. *PLoS One* 2014;9:e87315.
40. Marik PE, Stephenson E. The ability of procalcitonin, lactate, white blood cell count and neutrophil-lymphocyte count ratio to predict blood stream infection. Analysis of a large database. *J Crit Care* 2020;60:135-9.
41. Naffaa M, et al. Procalcitonin and interleukin 6 for predicting blood culture positivity in sepsis. *Am J Emerg Med* 2014;32:448-51.
42. Nishino T, et al. Evaluation of a proposed scoring system compared to procalcitonin levels and presepsin levels as predictors of a positive blood culture for bacteremia in acute cholangitis. *Gastroenterology* 2017;152:S500.
43. Oksuz L, et al. Procalcitonin and C-reactive protein in differentiating to contamination from bacteremia. *Braz J Microbiol* 2015;45:1415-21.
44. Oura H, et al. Diagnostic value of the serum lactate level as a predictor of bacteremia in patients with acute cholangitis. *J Hepatobiliary Pancreat Sci* 2023;30:1055-64.
45. Yañez Palma MC, et al. Utilidad de la región medial de la pro-adrenomodulina para la detección de bacteriemia verdadera en pacientes mayores atendidos en urgencias. *Rev Esp Quimioter* 2024;37:78-87.
46. Romualdo LG, et al. Diagnostic accuracy of presepsin (soluble CD14 subtype) for prediction of bacteremia in patients with systemic inflammatory response syndrome in the Emergency Department. *Clin Biochem* 2014;47:505-8.
47. Roy A, et al. Low procalcitonin can effectively rule out bacteremia. *Hosp Med* 2018; April 8-11;13. Orlando, FL. Abstract number 342.
48. Seo K, et al. Characteristics of serum endocan levels in infection. *PLoS One* 2015;10:e0123358.
49. Takahashi K, et al. Epidemiology, microbiology, and diagnosis of infection in diabetic ketoacidosis and hyperosmolar hyperglycemic syndrome. *Diabetes Res Clin Pract* 2024;212:111713.
50. Timbrook TT, et al. Procalcitonin and risk prediction for diagnosing bacteremia in hospitalized patients: A retrospective, national observational study. *Diagnostics (Basel)* 2023;13:3174.
51. Toreh C, et al. Procalcitonin levels predict bacteremia in patients with urosepsis. *Int J Urol* 2017;24:87.
52. Yurttutan Uyar N, et al. Sepsis biomarkers for early diagnosis of bacteremia in emergency department. *J Infect Dev Ctries* 2023;17:832-9.
53. Wiwatcharagoses K, Kingnakom A. Procalcitonin under Investigation as a means of detecting severe sepsis, septic shock and bacteremia at emergency department, Rajavithi hospital. *J Med Assoc Thai* 2016;99(Suppl 2):S63-8.

54. Xiao H, et al. Comparison among presepsin, procalcitonin, and C-Reactive protein in predicting blood culture positivity and pathogen in sepsis patients. *Shock* 2024;1(61):387-94.
55. Yamaguchi A, et al. Proportion of neutrophils in white blood cells as a useful marker for predicting bacteremic acute cholangitis. *Intern Med* 2023;62:2795-802.
56. Youkhana K, et al. The clinical utility of serum procalcitonin as a diagnostic assay for predicting bacteremia in patients with sepsis. *Chest* 2017;152(Suppl):A347.
57. Zhang CY, et al. Clinical value of alkaline phosphatase on the surface membrane of neutrophils for prediction of bacteremia in patients with systemic inflammatory response syndrome. *Diagn Microbiol Infect Dis* 2021;100:114105.
58. Zhang HY, et al. Predictive value of presepsin and acylcarnitines for severity and biliary drainage in acute cholangitis. *World J Gastroenterol* 2023;29:2502-14.
59. Lautenbach E, Localio R, Nachamkin I. Clinicians required very high sensitivity of a bacteremia prediction rule. *J Clin Epidemiol* 2004;57:1104-6.
60. Howie N, Gerstenmaier JF, Munro PT. Do peripheral blood cultures taken in the emergency department influence clinical management? *Emerg Med J* 2007;24:213-4.
61. Mountain D, Bailey PM, O'Brien D, Jelinek GA. Blood cultures ordered in the adult emergency department are rarely useful. *Eur J Emerg Med* 2006;13:76-9.
62. Meier MA, et al. Procalcitonin-guided antibiotic treatment in patients with positive blood cultures: A patient-level meta-analysis of randomized trials. *Clin Infect Dis* 2019;69:388-96.
63. Schuetz P, et al. Economic evaluation of procalcitonin-guided antibiotic therapy in acute respiratory infections: a US health system perspective. *Clin Chem Lab Med* 2015;53:583-92.



Chapter 3

Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model

Anna Kaal, Soufian Meziyerh, Nathalie van Burgel, Martijn Dane, Nikki Kolfschoten, Prashant Mahajan, Agustin Julián-Jiménez, Ewout Steyerberg, Cees van Nieuwkoop

ABSTRACT

Objectives

Blood cultures (BCs) are commonly ordered in emergency departments (EDs), while a minority yields a relevant pathogen. Diagnostic stewardship is needed to safely reduce unnecessary BCs. We aimed to develop and validate a bacteremia prediction model for ED patients, with specific focus on the benefit of incorporating procalcitonin.

Methods

We included adult patients with suspected bacteremia from a Dutch ED for a one-year period. We defined 23 candidate predictors for a “full model”, of which nine were used for an automatable “basic model”. Variations of both models with C-reactive protein and procalcitonin were constructed using LASSO regression, with bootstrapping for internal validation. External validation was done in an independent cohort of patients with confirmed infection from 71 Spanish EDs. We assessed discriminative performance using the C-statistic and calibration with calibration curves. Clinical usefulness was evaluated by sensitivity, specificity, saved BCs, and Net Benefit.

Results

Among 2111 patients in the derivation cohort (mean age 63 years, 46% male), 273 (13%) had bacteremia, versus 896 (20%) in the external cohort ($n = 4436$). Adding procalcitonin substantially improved performance for all models. The basic model with procalcitonin showed most promise, with a C-statistic of 0.87 (0.86–0.88) upon external validation. At a 5% risk threshold, it showed a sensitivity of 99% and could have saved 29% of BCs while only missing 10 out of 896 (1.1%) bacteremia patients.

Conclusions

Procalcitonin-based bacteremia prediction models can safely reduce unnecessary BCs at the ED. Further validation is needed across a broader range of healthcare settings.

INTRODUCTION

Prompt identification of bacteremia in patients with suspected infection is crucial due to potential high morbidity and mortality without timely treatment.¹ Currently, there is no satisfying rapid test to diagnose bacteremia in the emergency department (ED). According to sepsis guidelines, the standard is to obtain two sets of blood cultures (BCs).² Unfortunately, culture results are only available after 24–48 h. In the meantime, physicians often decide on the start of empirical antibiotics. The yield of ED BCs is disappointing, only 4–11% identify true pathogens, while 0.2–3% actually impact treatment decisions.^{3–12} Moreover, 2–8% of BCs are contaminated by skin flora, potentially resulting in longer hospital stay, increased costs and unnecessary exposure to antibiotics.^{3–8} Procalcitonin is a biomarker that has been suggested to be useful to identify patients at low risk of bacteremia, with sensitivity for bacteremia ranging from 76–99% depending on the cut-off level.^{13–16} Furthermore, clinical decision rules such as the Shapiro score have a sensitivity over 90%.^{17–21} It has been suggested that combining the Shapiro score with PCT could save 20% of BCs while identifying all bacteremia cases.¹⁵ However, biomarkers and decision rules have not found their way into routine clinical practice. We hypothesize that this is because simple models lack performance, while extensive decision rules might be too labor-intensive for busy EDs.²² Recently, multiple studies have developed promising automatable models that could be implemented in electronic health records (EHRs).^{23–26} Performance of such models might be improved by including PCT.

We aimed to develop an easy-to-use, automatable bacteremia prediction model to safely improve blood culture efficiency, preventing possible negative effects of false-positive BCs, while evaluating the value of incorporating procalcitonin. Moreover, we validated the potential clinical usefulness of such a model in an unselected and heterogeneous group of patients presenting with suspected infection at the ED.

MATERIALS AND METHODS

Study design and setting

We included consecutive patients aged 18 years and older with a BC taken between April 2019 and May 2020 during their ED visit at the Haga Teaching Hospital (a large hospital in The Hague, the Netherlands, with $\pm$ 60.000 ED visits annually). In the Dutch healthcare system, patients visit the ED if referred by their general practitioner or ambulance service. An automated PCT ordering procedure was installed for this prospective study. The decision to obtain a BC was left to the attending physician. PCT was measured using leftover blood, drawn for routine

chemistry concurrently with the BC. PCT was measured with an automated rapid sensitive assay (Elecsys Brahms PCT on Roche Cobas E601, Basel, Switzerland; measuring range: 0.02–100 ng/mL). The PCT result was unavailable for the attending physician. We only included the first visit in case of multiple ED visits during the study period. The TRIPOD guideline was followed for reporting on the development and validation of multivariable prediction models.²⁷

Data collection

Patient records were reviewed by one of five GCP-certified investigators. We collected data on socio-demographics, comorbidities as defined by the Charlson Comorbidity Index (CCI), medication use, and the first vital signs at the ED.²⁸ All numerical values were checked for impossible values. Additionally, we collected details on laboratory and microbiological results, and the clinical diagnosis in the ED.

Outcomes

The primary outcome was bacteremia. Patients were considered to have bacteremia if at least one BC obtained within 48 h of ED presentation yielded a pathogen that was not deemed a contaminant. For most cultures, this was assessed by the clinical microbiologist. In case of doubt, this was done by the hospital antibiotic stewardship team. Coagulase-negative staphylococci (CNS), *Bacillus* spp., *Corynebacterium* spp., *Propionibacterium* spp., Viridans group streptococci, and *Lactobacillus* spp. were considered contaminants unless associated with an intravascular catheter or device, suspected endocarditis, or multiple positive blood cultures with the same bacterium in a patient with a non-vascular prosthetic device (e.g. prosthetic joint). BCs were incubated using the BacTEC/ Alert automated BC system. Determination was done with gram stain and the Bruker Maldi-tof-MS Biotyper. Contaminated BCs were considered negative. Secondary outcomes were hospital admission, ICU-admission, length of hospitalization and 30-day mortality (see Supplement).

Definitions

Altered mental status was defined as a Glasgow coma scale < 15 or a note of confusion or changed behavior. Recent use of antibiotics was defined as use in the previous seven days. Diagnoses were divided into different suspected infection sites: Respiratory, Urogenital, Skin and soft tissue, Central nervous system, Abdominal, Intravascular, Bone and joint, Infection of unknown origin and No or other infection.²⁹ Two investigators (AK, ED) independently categorized the patients and in case of disagreement, an infectious disease specialist was consulted (CvN). Cohen's kappa was 0.87, indicating excellent agreement between assessors.

Ethics

This study was approved by the Medical Ethics Committee and the Institutional Scientific Review Board (protocol T18–040). The need for informed consent was waived given the observational nature without impact upon patient's management. This study was registered at the National Trial Register: <https://onderzoekmetmensen.nl/nl/trial/23345>.

Validation dataset

External validation was performed in an independent Spanish dataset from the INFURG-SEMES study group, including patients with suspected infection and BC testing from 71 EDs.²¹ The inclusion criteria and collection of variables were comparable to the Dutch cohort, apart from slightly different definitions in use of immunosuppressants, use of antibiotics, altered mental status and site of infection. This validation set differed from the development as ultimately, only patients with confirmed infection were included. See the original article for an extensive description of the methodology used in the Spanish dataset.²¹ As breathing frequency was dichotomized in the Spanish dataset, we used multiple imputation of the continuous breathing frequency with an indicator variable for missingness.

Statistical analysis

Continuous variables are presented as mean and standard deviation (SD) or median and interquartile range (IQR). Categorical data are summarized as counts and percentages. Comparisons were done using χ^2 and Mann-Whitney-U tests, as appropriate. We predefined 23 candidate predictors based on previous literature and expert knowledge (Table 2). We balanced the number of candidate predictors to the number of bacteremia cases aiming for a minimum of 10 events per predictor to prevent over-fitting, including all patients for a one-year period to cover all seasons.³⁰ All candidate predictors were used for a 'full' model, whereas only readily automatable predictors were used in a 'basic' model. Prediction models were fitted without biomarkers, with either CRP or PCT alone, and with both combined. All numerical predictors were used as a continuous variable. We considered linear, logistic or spline transformations of candidate predictors as indicated upon visual inspection. Variables with extreme values were winsorized towards the 1st and 99th percentile.³⁰ There was a maximum of 9% missing data per candidate predictor, which were multiple imputed using chained equations.³¹ Results were compared to those from complete case analysis (eTable 5). We created a stacked dataset of 10 imputed datasets, with weights of 1/10 per patient. Regression analysis with the Least Absolute Shrinkage and Selection Operator (LASSO) was used to

define the models. Odds ratios and intercepts are given without 95% confidence (CI) interval since penalization with LASSO leads to biased estimates that are optimal for prediction (eTable 6, eTable 7). Predictive performance was assessed by discrimination and calibration. Discriminative ability was assessed by the C-statistic, which is the area under the ROC curve. Calibration was evaluated by the calibration intercept, slope and graphically by plotting observed outcomes versus predicted risk. Internal validation was conducted with a 200 bootstrap resampling procedure, where the LASSO procedure was repeated in each bootstrap sample to estimate the optimism-corrected C-statistic.³⁰ Predictions were calculated using the predict function from the R package 'stats'. Final models were constructed in the combined development and validation dataset without predictors with coefficients close to zero. For these models, we repeated development and internal validation using a weighted LASSO regression with dataset (Spanish vs Dutch) as an additional predictor and using equal weighting for each dataset. As a safety assessment, we performed a subgroup analysis for immunocompromised patients. P-values of < 0.05 were considered statistically significant. Analyses were performed using RStudio version 2023.03.0.

Potential clinical usefulness

We calculated sensitivity, specificity, positive predictive value, negative predictive value, and the corresponding saved BCs and missed positive BCs. We report outcomes at a risk threshold of 5% for comparability to previous literature and for thresholds from 1% to 10% (eTable 3).²⁶ For PCT, we considered preset cut-offs of < 0.05 ng/mL, < 0.10 ng/mL, < 0.25 ng/mL and < 0.50 ng/mL. Clinical usefulness was summarized as Net Benefit (NB), expressed as net interventions avoided per 100 patients, also known as NBOpt-out.^{32–34} NB quantifies the potential impact on patient outcomes by decision-making based on a prediction model compared to reference strategies of testing all or none of the patients (see Supplement). We constructed decision curves to show NB between 1% and 25% thresholds. We consider risk thresholds of 1% to 10% most clinically relevant, and above 25% clinically implausible since any clinician would order a BC if the risk of bacteremia would be over 25%.

RESULTS

Of 2550 ED visits in the Dutch cohort, 218 were excluded because PCT was not correctly measured and 221 were patients with multiple ED visits (eFig. 1). Of 2111 unique patients, 273 were diagnosed with bacteremia (13%) and 80 patients had contaminated BCs (3.8%). A total of 2980 BCs were taken from these 2111 patients

of which 404 yielded a true pathogen (14%) and 84 (2.8%) were contaminated. For detailed culture results, see eTable 1.

The mean age in the Dutch dataset was 63 years and 54% (1147/2111) was male (Table 1). Bacteremia patients had more cardiovascular disease (146/273, 54% versus 668/1838, 36%), more diabetes mellitus (78/273, 29% versus 388/1838, 21%) and more often a history of malignancy (78/ 273, 29% versus 394/1838, 21%) (eTable 2). They suffered from more comorbidities (median CCI 4 versus 3). Furthermore, vital parameters reflected more severe disease for those with bacteremia (e.g. lower blood pressure, higher temperature, heart rate and breathing frequency).

Median PCT in bacteremia patients was 1.3 (IQR [0.43–9.3]) versus 0.18 (IQR [0.08–0.60]) for those without bacteremia ($p < 0.001$). Clinical outcomes were worse for bacteremia patients compared to those without bacteremia (See Supplement). The 4436 Spanish patients had similar demographics (mean age 67 years; 2646/4436, 60% male) and less comorbidity (CCI median of 2 [1–4], Table 1). Mean systolic blood pressure was lower and inflammatory markers were higher. Bacteremia was more frequent (896/4436, 20%), and more BCs were contaminated (227/4436, 5.1%).

Table 1: Baseline characteristics for the Dutch and Spanish dataset

Characteristics	Netherlands No Bacteremia n = 1838	Netherlands Bacteremia n = 273	Spain No Bacteremia n = 3540	Spain Bacteremia n = 896
Demographics and medical history				
Age, years, (mean (SD))	61.9 (19.3)	69.9 (15.2)	66.5 (19.0)	70.3 (16.0)
Sex (Male)	992 (54.0)	155 (56.8)	2099 (59.3)	547 (61.0)
Sex (Female)	847 (46.1)	117 (42.9)	1441 (40.7)	349 (39.0)
Recent use of antibiotics*	475 (25.8)	36 (13.2)	566 (16.0)	106 (11.8)
Use of immunosuppressants	220 (12.0)	32 (11.7)	256 (7.2)	75 (8.4)
Charlson Comorbidity Index (median [IQR])	3.0 [0.0, 5.0]	4.0 [2.0, 6.0]	2.0 [1.0, 4.0]	3.0 [1.0, 5.0]
Vital functions in ED				
Temperature, °C, (mean (SD))	38.3 (1.1)	38.8 (1.1)	37.8 (1.0)	38.1 (1.1)
Glasgow coma scale (median [IQR])	15.0 [15.0, 15.0]	15.0 [15.0, 15.0]	15.0 [15.0, 15.0]	15.0 [14.0, 15.0]
Heart frequency, beats/min, (mean (SD))	100.0 (21.0)	106.3 (23.9)	99.0 (19.9)	103.6 (22.5)
Systolic blood pressure, mmHg, (mean (SD))	133.8 (27.1)	123.8 (32.3)	124.3 (27.0)	114.4 (27.6)
Breathing frequency > 22 breaths/min	536 (32.3)	113 (44.5)	1115 (34.5)	523 (58.5)
Breathing frequency, breaths/min, (median [IQR])	20.0 [16.0, 24.0]	20.5 [16.0, 28.0]		
Saturation, %, (median [IQR])	96.0 [94.0, 98.0]	95.0 [93.0, 97.0]		
Laboratory values in ED				
C-reactive protein, mg/L, (median [IQR])	68.0 [23.0, 159.5]	122.0 [38.0, 238.0]	110.0 [50.0, 211.0]	159.0 [81.2, 265.5]
Procalcitonin, µg/L, (median [IQR])	0.2 [0.1, 0.6]	1.3 [0.4, 9.2]	0.2 [0.1, 0.5]	3.1 [1.1, 10.7]
White blood cell count, ×10 ⁹ /L, (median [IQR])	11.1 [7.8, 15.3]	12.6 [8.9, 17.3]	11.6 [8.0, 16.3]	15.1 [10.3, 19.3]
Creatinine, µmol/L, (median [IQR])	80.5 [65.0, 106.0]	100.0 [78.0, 140.5]		
Suspected site of infection*				
Respiratory	751 (40.9)	43 (15.8)	1783 (50.4)	187 (20.9)

Characteristics	Netherlands No Bacteremia n = 1838	Netherlands Bacteremia n = 273	Spain No Bacteremia n = 3540	Spain Bacteremia n = 896
Urogenital	263 (14.3)	96 (35.2)	894 (25.3)	382 (42.6)
Abdominal	182 (9.9)	43 (15.8)	355 (10.0)	170 (19.0)
Skin and soft tissue	145 (7.9)	24 (8.8)	163 (4.6)	45 (5.0)
Central nervous system	15 (0.8)	7 (2.6)	16 (0.5)	18 (2.0)
Unknown site of infection	113 (6.1)	20 (7.3)	221 (6.2)	60 (6.7)
Other infection ^b	98 (5.3)	4 (1.5)	108 (3.1)	34 (3.8)
Bone and joint	22 (1.2)	16 (5.9)		
Intravascular	12 (0.7)	11 (4.0)		
No infection	239 (13.0)	9 (3.3)		

Table 1. Baseline characteristics for the Dutch (derivation) and Spanish (validation) dataset. Data are presented as number (percentage) of patients unless otherwise indicated. SD = Standard Deviation, IQR = Interquartile Range, ED = Emergency Department. Missing in Dutch data: (n (%)): Temperature 3 (0.1), Glasgow coma scale 27 (1.3), Heart frequency 21 (1.0), Systolic blood pressure 22 (1.0), Breathing frequency 197 (9.3), Saturation 23 (1.1), Creatinine 7 (0.3). Missing in Spanish data (n (%)): Temperature 14 (0.3), Glasgow coma scale 91 (2.1), Heart frequency 67 (1.5), Systolic blood pressure 30 (0.7), Breathing frequency >22 307 (6.9), C-reactive protein 362 (8.2), Procalcitonin 140 (3.2), White blood cell count 15 (0.3). ^a Recent use of antibiotics in the previous seven days. ^b Two patients were diagnosed with multiple infections. ^c Composite category of multiple infections such as viral infections or ear-nose-throat infections.

Table 2: Bacteremia prediction models with odds ratios per predictor

Candidate predictors	Basic without biomarkers	Basic with CRP	Basic with PCT	Basic with CRP and PCT	Full without biomarkers	Full with CRP	Full with PCT	Full with CRP and PCT	Only PCT
Age	1.4	1.4	1.3	1.3	1.4	1.4	1.3	1.3	
Gender									
SBP	1.9	1.9	1.3	1.2	1.9	1.8	1.3	1.2	
BFR		0.998	0.95	0.95	1.2	1.2	1.1	1.1	
HFR	1.2	1.2	1.2	1.2	1.4	1.4	1.3	1.3	
Temperature	2.2	2.2	2.1	2.1	1.9	2.0	2.0	2.0	
GCS	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	
CRP		1.1		0.95		1.1		0.97	
PCT			1.4	1.4			1.4	1.4	1.4
Antibiotics					0.4	0.4	0.4	0.4	
Immuno						1.03	1.01	1.01	
AMS					0.9	0.9			
Catheter					3.5	3.5	4.4	4.4	
Chills					1.6	1.6	1.4	1.4	
Site of infection ^a					0.9-19	0.7-14	0.7-17	0.7-19	

Table 2. Overview of all models derived from Dutch derivation dataset with corresponding odds ratio per predictor. Predictors in red were not selected for the final model in the LASSO procedure. Grey variables were not considered. CRP = C-reactive protein, PCT = procalcitonin, SBP = systolic blood pressure, BFR = breathing frequency, HFR = heart frequency, GCS = glasgow coma scale, Antibiotics = Use of antibiotics in previous 7 days, Immuno = use of immunosuppressants, AMS = altered mental status, Catheter = indwelling vascular catheter, Chills = reported chills during ED visit. Possible infection sites were: No infection/non-bacterial infection, respiratory, skin and soft tissue, unknown, abdominal, urogenital, central nervous system, bone and joint, intravascular. The odds ratios reflect the following change in values: Age = per 10 years increase, SBP = per 20 mmHg decrease, BFR = increase of breathing frequency of 5/min, HFR = per increase in heartrate of 30 beats/min, GCS for each point decrease, CRP and PCT = for a doubled value. ^aFor odds ratios per infection site see Supplemental material.

Table 3: Performance of all models upon external validation in the Spanish dataset.

Model	Performance		Calibration		Outcomes for a 5% risk threshold				
	Discrimination		Intercept (95% CI)		Slope (95% CI)	Saved BC (%)	Missed TP (%)	Sens (%)	NPV (%)
	Internal (95% CI)	External (95% CI)	Intercept (95% CI)	Slope (95% CI)					
Basic without biomarkers	0.72 (0.69 - 0.75)	0.66 (0.64-0.68)	0.53 (0.45-0.60)	0.74 (0.64-0.85)	0.74 (0.64-0.85)	10.4	31 (3.5)	96.5	93.3
Basic with CRP	0.73 (0.70 - 0.76)	0.67 (0.65-0.69)	0.45 (0.37-0.53)	0.77 (0.67-0.87)	0.77 (0.67-0.87)	10.4	26 (2.9)	97.1	94.4
Basic with PCT	0.81 (0.78 - 0.84)	0.87 (0.86-0.88)	0.59 (0.51-0.67)	1.44 (1.34-1.53)	1.44 (1.34-1.53)	28.5	10 (1.1)	98.9	99.2
Basic with CRP and PCT	0.81 (0.78 - 0.84)	0.88 (0.86-0.89)	0.62 (0.54-0.70)	1.42 (1.33-1.52)	1.42 (1.33-1.52)	29.8	11 (1.2)	98.9	99.2
Full without biomarkers	0.82 (0.80 - 0.85)	0.76 (0.74-0.77)	0.27 (0.19-0.35)	0.81 (0.74-0.88)	0.81 (0.74-0.88)	25.7	63 (7.0)	93.0	94.5
Full with CRP	0.83 (0.80 - 0.85)	0.77 (0.75-0.78)	0.20 (0.12-0.28)	0.83 (0.76-0.90)	0.83 (0.76-0.90)	25.3	52 (5.8)	94.2	95.4
Full with PCT	0.86 (0.84 - 0.88)	0.88 (0.86-0.89)	0.41 (0.33-0.50)	1.20 (1.12-1.28)	1.20 (1.12-1.28)	37.4	25 (2.8)	97.2	98.5
Full with CRP and PCT	0.86 (0.84 - 0.88)	0.88 (0.87-0.89)	0.43 (0.35-0.52)	1.20 (1.12-1.28)	1.20 (1.12-1.28)	38.1	25 (2.8)	97.2	98.5
PCT only	0.79 (0.76 - 0.82)	0.90 (0.89-0.91)	0.46 (0.38-0.54)	1.67 (1.57-1.78)	1.67 (1.57-1.78)	14.5	3 (0.3)	99.7	99.5

Table 3. Performance of all models upon external validation in the Spanish dataset. CRP = C-reactive protein, PCT = procalcitonin, CI = confidence interval, BC = blood cultures, TP = True positive. Outcomes are given at patient-level. For example: applying the basic model without biomarkers would have saved blood cultures in 457 out of 4436 (10.3%) patients, but it would have missed 3.5% of all 896 true positive blood cultures.

Prediction models, performance and clinical usefulness

Various prediction models were constructed (Table 2). Upon development, the basic models including PCT showed the highest discriminative performance with C-statistics of 0.81 (0.78–0.84) (Table 3). Adding clinical information increased the C-statistic beyond those of all basic models. Again, including PCT in the full models led to a substantial increase in C-statistic. Upon external validation, the effect of adding PCT to the basic models led to even higher C-statistics of 0.87 (0.86–0.88). In contrast with the Dutch dataset, adding clinical variables did not show additional benefit over PCT. Upon external validation, the basic model with PCT saved BCs in 29% (1266/4366) of patients while only missing 10 out of 896 (1.1%) bacteremia patients. Adding CRP would have missed one more bacteremia while saving another 1.3% of BCs. A model with procalcitonin only had a C-statistic of 0.79 (0.76–0.82) at internal validation and 0.90 (0.89–0.91) at external validation. This model saved 15% (643/4436) of BCs and missed 3 out of 896 bacteremia patients. Prediction models had limited impact on BC ordering if a risk threshold of 1% was applied, and more impact for a 10% threshold (eTable 3). Overall, calibration showed systemic under-estimation of risk in the Spanish cohort, with intercepts of all models significantly different from 0 (Fig. 1). The calibration plots of models including procalcitonin showed a S-shape with overestimation of risks in the lower risk ranges and underestimation in the higher risk ranges, reflected in slopes significantly different from 1 (Fig. 1). A set of final updated models can be found in eTable 7. Finally, both the original and updated models that included PCT provided a net reduction of BCs across the vast majority of risk thresholds, specifically in the relevant range of thresholds between 1% and 10%. (Fig. 2).

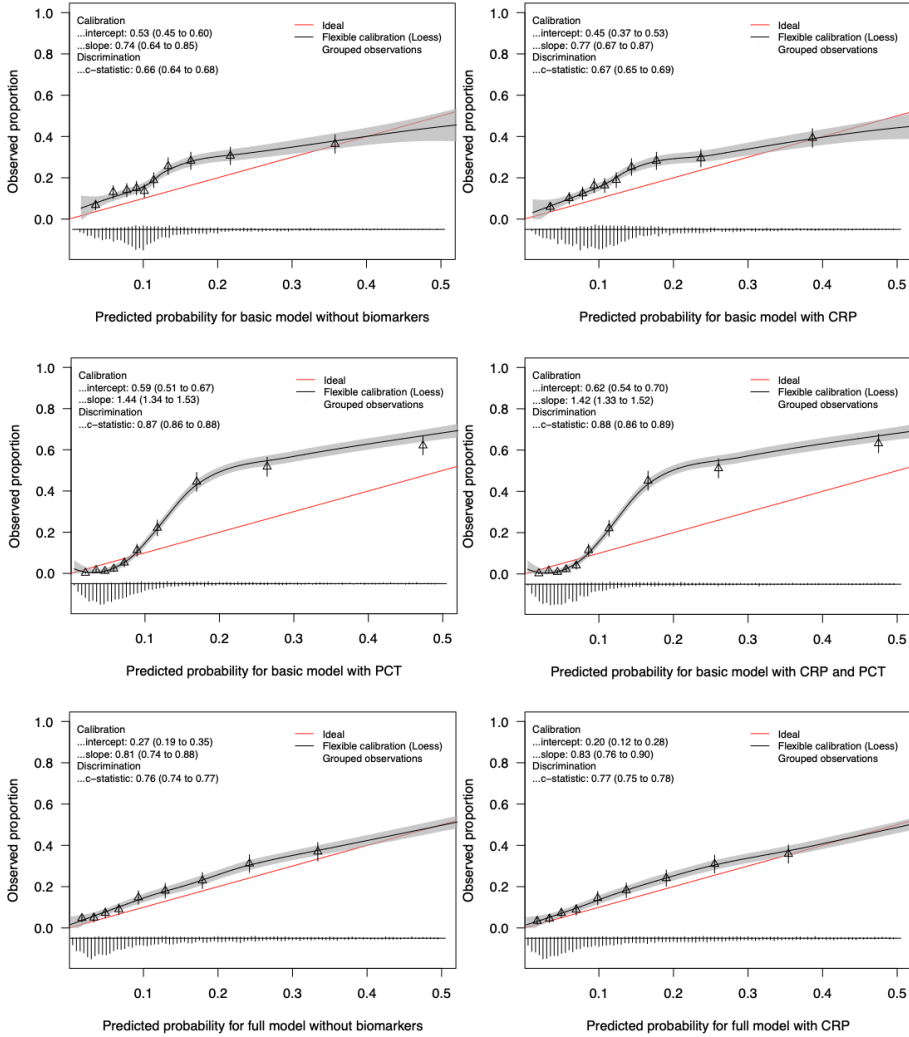
Missed bacteremia

Applying a 5% risk threshold in the basic with PCT model in the Dutch data would have withheld BCs in 733/2111 (35%) of patients. This reduction would imply missing 12 bacteremia patients. A chart review suggested a probable negative impact for 2 out of 12 patients and one case with possible negative impact. For the other 9 patients, no negative impact was expected (eTable 8).

Subgroup analysis

Among 331 immunocompromised Spanish patients, 75 (23%) had bacteremia. Using the basic model + PCT at the 5% risk threshold, none of those would have been missed while avoiding 27% (89/331) of BCs. The C-statistic was 0.90 (0.86–0.93), and the underestimation of risk was similar to the full validation set (calibration intercept 0.71, 0.42–1.00).

Figure 1: Calibration plots of all models applied in the Spanish dataset



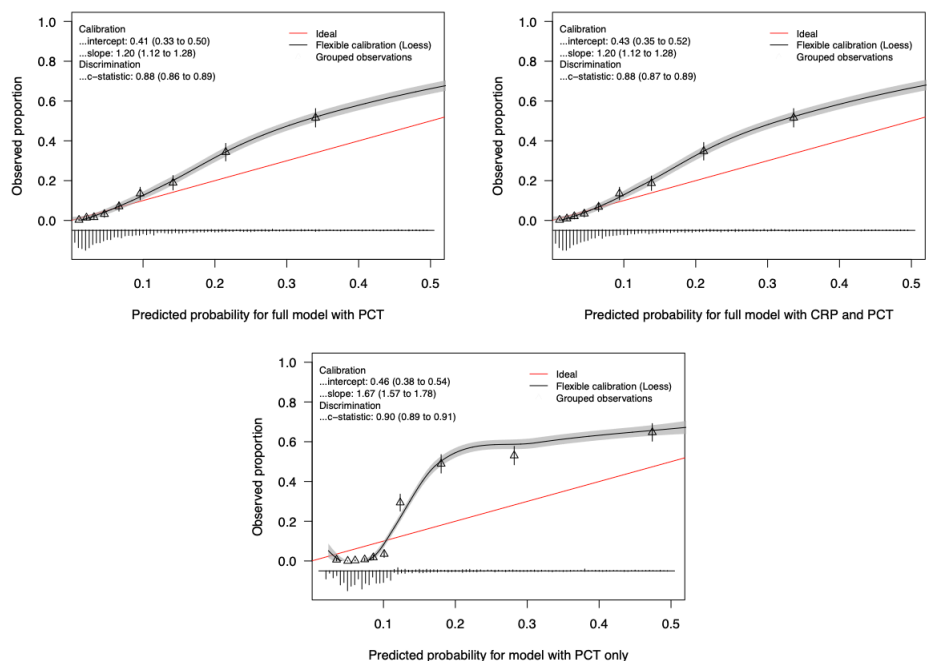


Figure 1. Calibration plots of all models applied in the Spanish dataset. X-axis represents predicted risks, Y-axis is the observed risks in the Spanish dataset. CRP = C-reactive protein, PCT = procalcitonin.

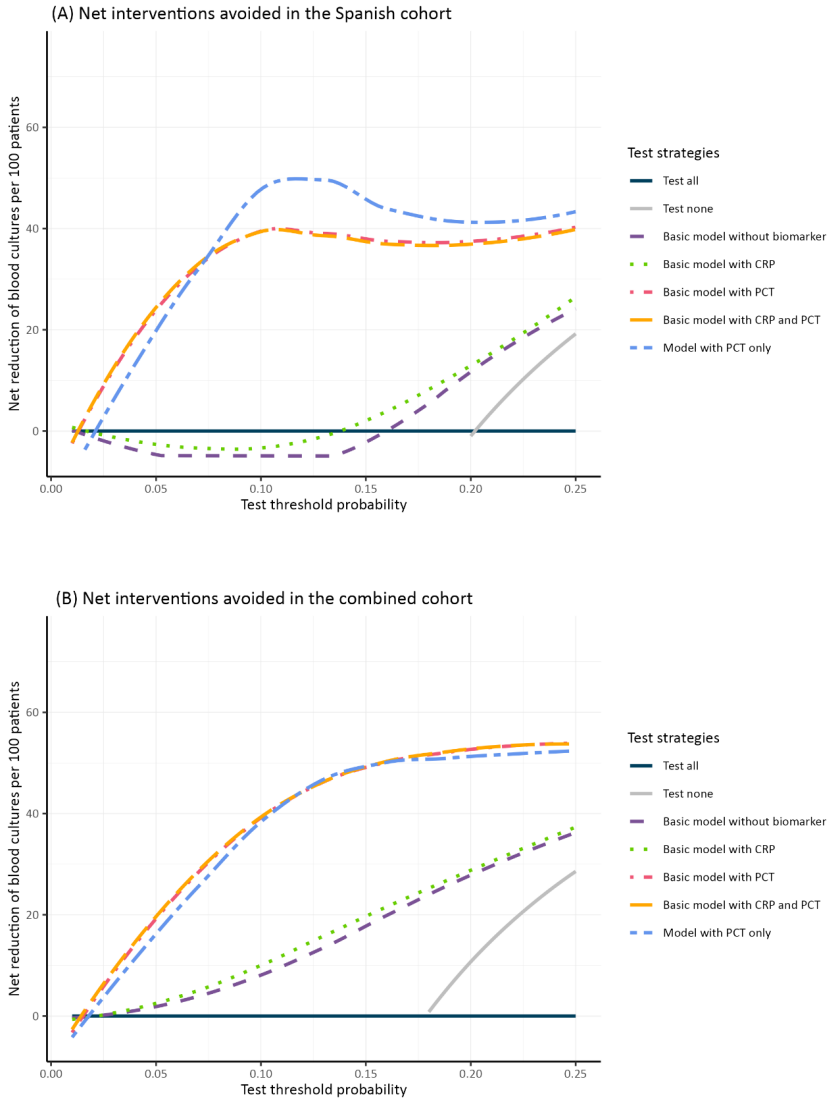
Figure 2: Net reduction of blood cultures


Figure 2. Decision curve analysis showing the expected net reduction of blood cultures per patient for models in the Spanish cohort (A) and in the combined cohorts (B). The net reduction of blood cultures relative to testing everyone (“Test all”) is shown. The x-axis displays the possible risk thresholds for bacteremia on which to decide to withhold blood cultures. For instance: one physician might accept a maximum bacteremia risk of 5%, taking cultures in most patients. The net reduction of blood cultures, or Net BenefitOpt-out, can be calculated by $TNRR \cdot (1-p) - p \cdot (1-R/R) \cdot FNRR$, in which TNR is the true negative rate, R is the risk threshold, p is the disease prevalence and FNR is the false negative rate. At a 5% threshold, the net benefit of the basic model with PCT is 0.24, calculated as $((1256/3540) \cdot 0.798) - ((10/896) \cdot 0.202 \cdot (0.95/0.05))$. See the supplemental material for more information about Net Benefit.

DISCUSSION

We found a substantial incremental value of PCT in multiple bacteremia prediction models, enabling a safe reduction in BC ordering. Five of our nine bacteremia prediction models were designed to be fully automatable, allowing for easy validation and implementation in EHRs. Considering both predictive performance and applicability, the basic model with PCT seems most promising. It could potentially save 29% of BCs, while only missing 1% of bacteremia patients.

While we investigated multiple models, we focused on the basic with PCT model since including PCT led to substantially improved predictive performance. This is in line with previous research which showed an increase in the C-statistic of the Shapiro score from 0.73 to 0.80 by adding PCT.¹⁵ In contrast, Lien et al. showed no added benefit of PCT to a machine learning model based on a complete blood count.³⁵ The added benefit of detailed clinical information was limited in both cohorts, while the manual data-entry required for such information would be a barrier for implementation at the ED. Although the basic model with both PCT and CRP as biomarkers was able to save slightly more BCs, the incremental value might be considered too limited to justify testing CRP for all patients. PCT had substantial value on its own already, particularly in the Spanish data set.

Interpretation

A previously developed machine learning model demonstrated a C-statistic of 0.81 (0.78–0.83) in a test cohort and, in a prospective validation, saved 30% of BCs while missing 6.6% of bacteremia patients at a 5% risk threshold.²⁶ This reflects a sensitivity of 93%, while our basic with PCT model had a sensitivity of 99%, which is what clinicians appreciate for a bacteremia prediction model.²² Our models have the advantage of being transparent with respect to the predictors and their weights, unlike machine learning models in which variables are unavailable to the clinician, a 'black-box'. Second, choosing candidate predictors based on known physiological processes is expected to lead to more robust models with validity in different settings.³⁶ Finally, being able to provide the regression equation also enables more straightforward validation and implementation in new settings.³⁷

The discriminative performance of our models was higher in the Spanish cohort than at model development, but calibration showed systematic underestimation. This might relate to the different patient selection process for the Spanish dataset, as only patients with confirmed infection were included. Our models were unable to fully capture this difference in average bacteremia risk. Furthermore, the predictive value of PCT was very strong in the Spanish dataset, which explains the improved performance for all models containing PCT upon external validation.

We do not have a clear explanation for this difference in diagnostic strength of PCT. The miscalibration motivated updating of the prediction models based on both datasets (eTable 6). These updated models require further external validation before considering clinical implementation. By following an iterative process of model updating and validation, we aim to create a broadly applicable bacteremia prediction model.^{38,39}

In times where diagnostic stewardship becomes increasingly important, a bacteremia prediction model like ours meets a pressing need. However, no model will be perfect, and the use of such a model will inevitably miss some bacteremia patients. Although clinicians are hesitant to miss bacteremia, there is extensive literature on the lack of treatment consequences in the majority of positive BCs.^{3,9,10} Indeed, upon clinical review, missing bacteremia in patients with a predicted low risk was generally deemed clinically irrelevant, except in those with suspected endocarditis (eTable 8). We therefore do not recommend using our models in patients with a clear indication for BCs, such as suspected endocarditis or central-line associated bacteremia. Notably, the model performed well in the Spanish immunocompromised patients, missing none of 75 bacteremia cases at a threshold of 5%.

Societal value

Apart from the clinical usefulness of our models, they also have potential societal value through cost reduction. Applying the basic with PCT model in the Spanish dataset would save 571 negative BCs per 1000 patients, assuming two BCs per patient. Additionally, it would prevent 18 false positive BCs per 1000 patients. Assuming 96 to 128 USD per negative BC, this would potentially save k\$55 to k\$73 per 1000 patients.⁷ At estimated costs of \$2923 to \$5812, avoiding the false positive BCs would save an additional \$53.000 to \$105.000, leading to a total saving of \$108.000 and \$178.000 per 1000 patients. Testing PCT for 1000 patients is estimated to cost \$51.000 to \$96.000.^{40,41} This would result in a potential net cost-saving of \$12.000 to \$127.000 per 1000 patients. We recognize that the costs of BCs and PCT testing differ across the world, and PCT testing can be costly, which will influence the potential cost-saving. Additionally, it is difficult to quantify the costs of missing a bacteremia. In the future, a more extensive cost-benefit analysis should be performed to quantify societal value.

Implementation

There are two main things to be considered before implementing these models in clinical practice. First, we believe that another validation study of the updated models is essential to increase our understanding of the observed geographical

differences. Second, an implementation study should evaluate whether the models indeed lead to a safe reduction of BCs, while exploring how clinicians use such models and what risk of bacteremia they are willing to accept.

We believe that our model should only be used in patients where the clinician sees an indication for BCs. This would prevent using the model in a broader range of patients, which could potentially lead to an unwanted increase of BC testing. It is important to realize that we only included patients with a clinical indication for BCs, and do not know the model performance in patients where the clinician would normally not have taken BCs. Finally, it is important to realize that the potential cost-saving will be different per setting, mainly depending on the costs of BC and PCT testing. In places where PCT testing is not already implemented, high costs of PCT might be a barrier for implementation. However, we believe that especially at EDs already using PCT, the model has the potential to lead to significant cost reduction and can prevent patient harm related to false-positive BCs.

Strengths and limitations

The main strength of this study is the development and external validation of our models in two large international datasets with high-quality data and minimal missing values. While some definitions differed slightly between the development and validation cohorts, we consider the data a good reflection of clinical practice across different geographic locations as clinicians will interpret variables slightly different across settings.³⁹ Furthermore, we applied state-of-the-art statistical methods to develop robust prediction models. This robustness is underlined by the fact that the models performed well even though the validation cohort consisted only of confirmed infections, in contrast to suspected infections in the development data. A limitation is that we observed some unexplained geographic variation in statistical performance (discrimination and calibration) and clinical usefulness (Net Benefit). This motivates validating the updated models in other geographical settings. Finally, we emphasized clinical interpretability and applicability, which will make implementation in clinical practice more feasible.

CONCLUSIONS

PCT should be considered in any bacteremia prediction model because of considerably improved performance beyond clinical information alone. PCT-based bacteremia prediction models can potentially safely reduce a large number of unnecessary BCs at the ED, motivating further validation across a range of geographical locations and healthcare settings.

Funding

This project was supported by an unrestricted grant from the Haga Scientific Research Fund. This did not influence the study design, the collection and analysis of data, the writing of the manuscript or the decision to submit for publication.

CRedit authorship contribution statement

AK: Conceptualization, Methodology, Data curation, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing. **SM:** Conceptualization, Methodology, Data curation, Visualization, Writing – review & editing, Funding acquisition. **NvB:** Resources, Writing – review & editing. **MD:** Resources, Writing – review & editing. **NK:** Writing – review & editing. **PM:** Writing – review & editing. **AJ-J:** Writing – review & editing, Data curation. **ES:** Conceptualization, Methodology, Formal analysis, Validation, Visualization, Writing – review & editing, Supervision. **CvN:** Conceptualization, Resources, Methodology, Visualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used OpenAI in order to reduce word count. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content for publication.

Data availability

De-identified, individual patient data supporting the findings of this study are available from the corresponding author upon reasonable request, available from publication without end date. R Markdown files showing the analysis for one of the models are available in the Supplementary material.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: ES received royalties from Springer Verlag for his book “Clinical Prediction Models. AJJ declares to have received payments for presentations and lectures for Roche Diagnostics, Thermo Fisher Scientific, B.R.A.H.M.S. AG, Biomerieux and Virogates. All other authors reported no competing interests.

Acknowledgments

We want to thank Dagmar Brouwer, Dide van de Fliert, Nicoline Meijs, Lars Snel and Emma Dregmans for their help with retrieving all clinical information from the electronic patient records. Furthermore, we want to thank the INFURG-SEMES study group for sharing their data.

SUPPLEMENTAL MATERIAL

3



REFERENCES

1. Jiménez-Aguilar P, López-Cortés LE, Rodríguez-Baño J. *Impact of infectious diseases consultation on the outcome of patients with bacteraemia. Ther Adv Infect Dis* 2019 Dec 6;**6**:2049936119893576. <https://doi.org/10.1177/2049936119893576>. PMID: 31839942; PMCID: PMC6900613.
2. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. *Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock. Intensive Care Med* 2016;**43**(3):304–77. <https://doi.org/10.1007/s00134-017-4683-6>. Epub 2017 Jan 18. PMID: 28101605.
3. Long B, Koefman A. *Best clinical practice: blood culture utility in the emergency department. J Emerg Med* 2016 Nov;**51**(5):529–39. <https://doi.org/10.1016/j.jemermed.2016.07.003>. Epub 2016 Sep 14. PMID: 27639424.
4. Coburn B, Morris AM, Tomlinson G, Detsky AS. *Does this adult patient with suspected bacteremia require blood cultures? JAMA* 2012 Aug 1;**308**(5):502–11. <https://doi.org/10.1001/jama.2012.8262>
5. Nannan Panday RS, Wang S, van de Ven PM, Hekker TAM, Alam N, Nanayakkara PWB. *Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. PLoS One* 2019 Mar 21;**14**(3):e0214052. <https://doi.org/10.1371/journal.pone.0214052>. PMID: 30897186; PMCID: PMC6428292.
6. Dargère S, Cormier H, Verdon R. *Contaminants in blood cultures: importance, implications, interpretation and prevention. Clin Microbiol Infect* 2018 Sep;**24**(9):964–9. <https://doi.org/10.1016/j.cmi.2018.03.030>. Epub 2018 Apr 3. PMID: 29621616.
7. Dempsey C, Skoglund E, Muldrew KL, Garey KW. *Economic health care costs of blood culture contamination: a systematic review. Am J Infect Control* 2019 Aug;**47**(8):963–7. <https://doi.org/10.1016/j.ajic.2018.12.020>. Epub 2019 Feb 20. PMID: 30795840.
8. Bates DW, Goldman L, Lee TH. *Contaminant blood cultures and resource utilization. The true consequences of false-positive results. JAMA* 1991 Jan 16;**265**(3):365–9.
9. Howie N, Gerstenmaier JF, Munro PT. *Do peripheral blood cultures taken in the emergency department influence clinical management? Emerg Med J* 2007 Mar;**24**(3):213–4. <https://doi.org/10.1136/emj.2006.039875>. PMID: 17351231; PMCID: PMC2660034.
10. Mountain D, Bailey PM, O'Brien D, Jelinek GA. *Blood cultures ordered in the adult emergency department are rarely useful. Eur J Emerg Med* 2006 Apr;**13**(2):76–9. <https://doi.org/10.1097/01.mej.0000188231.45109.ec>. PMID: 16525233.
11. Fabre V, Milstone AM, Keller SC, Carroll KC, Cosgrove SE. *Prescribers' knowledge, attitudes and perceptions about blood culturing practices for adult hospitalized patients: a call for action. Infect Control Hosp Epidemiol* 2018 Nov;**39**(11):1394–6. <https://doi.org/10.1017/ice.2018.224>. Epub 2018 Sep 18. PMID: 30226121.
12. Dräger S, Giehl C, Søgaard KK, Egli A, de Roche M, Huber LC. *Do we need blood culture stewardship programs? A quality control study and survey to assess the appropriateness of blood culture collection and the knowledge and attitudes among physicians in Swiss hospitals. Eur J Intern Med* 2022;**103**:50–6. <https://doi.org/10.1016/j.ejim.2022.04.028>. Epub 2022 Jun 14. PMID: 35715280.

13. Hoeboer SH, van der Geest PJ, Nieboer D, Groeneveld AB. *The diagnostic accuracy of procalcitonin for bacteraemia: a systematic review and meta-analysis. Clin Microbiol Infect* 2015 May;**21**(5):474–81. <https://doi.org/10.1016/j.cmi.2014.12.026>. Epub 2015 Jan 14. PMID: 25726038.
14. van Nieuwkoop C, Bonten TN, van't Wout JW, Kuijper EJ, Groeneveld GH, Becker MJ, et al. *Procalcitonin reflects bacteremia and bacterial load in urosepsis syndrome: a prospective observational study. Crit Care* 2010;**14**(6):R206. <https://doi.org/10.1186/cc9328>. , Epub 2010 Nov 17. PMID: 21083886; PMCID: PMC3220019.
15. Laukemann S, Kasper N, Kulkarni P, Steiner D, Rast AC, Kutz A, et al. *Can we reduce negative blood cultures with clinical scores and blood markers? Results from an observational cohort study. Medicine* 2015;**94**(49):e2264. <https://doi.org/10.1097/MD.0000000000002264>
16. Goodlet KJ, Cameron EA, Nailor MD. *Low sensitivity of procalcitonin for bacteremia at an academic medical center: a cautionary tale for antimicrobial stewardship. Open Forum Infect Dis* 2020 Mar 16;**7**(4):ofaa096. <https://doi.org/10.1093/ofid/ofaa096>. PMID: 32322602; PMCID: PMC7162616.
17. Shapiro NI, Wolfe RE, Wright SB, Moore R, Bates DW. *Who needs a blood culture? A prospectively derived and validated prediction rule. J Emerg Med* 2008 Oct;**35**(3):255–64. <https://doi.org/10.1016/j.jemermed.2008.04.001>. Epub 2008 May 16. PMID: 18486413.
18. Bates DW, Cook EF, Goldman L, Lee TH. *Predicting bacteremia in hospitalized patients. A prospectively validated model. Ann Intern Med* 1990 Oct 1;**113**(7):495–500. <https://doi.org/10.7326/0003-4819-113-7-495>. PMID: 2393205.
19. Eliakim-Raz N, Bates DW, Leibovici L. *Predicting bacteraemia in validated models—a systematic review. Clin Microbiol Infect* 2015 Apr;**21**(4):295–301. <https://doi.org/10.1016/j.cmi.2015.01.023>. Epub 2015 Feb 10. PMID: 25677625.
20. Julián-Jiménez A, García-Lamberechts EJ, González Del Castillo J, Navarro Bustos C, Llopis-Roca F, Martínez-Ortiz de Zarate M, et al. *Validation of a predictive model for bacteraemia (MPB5-Toledo) in the patients seen in emergency departments due to infections. Enferm Infecc Microbiol Clin* 2022;**40**(3):102–12. <https://doi.org/10.1016/j.jimce.2021.12.006>. Epub 2022 Jan 4. PMID: 34992000.
21. Julián-Jiménez A, González Del Castillo J, García-Lamberechts EJ, Huarte Sanz I, Navarro Bustos C, Rubio Díaz R, et al. *A bacteraemia risk prediction model: development and validation in an emergency medicine population. Infection* 2022;**50**(1):203–21. <https://doi.org/10.1007/s15010-021-01686-7>. Epub 2021 Sep 6. PMID: 34487306.
22. Lautenbach E, Localio R, Nachamkin I. *Clinicians required very high sensitivity of a bacteremia prediction rule. J Clin Epidemiol* 2004 Oct;**57**(10):1104–6. <https://doi.org/10.1016/j.jclinepi.2004.03.003>. PMID: 15528062.
23. Ratzinger F, Dedeyan M, Rammerstorfer M, Perkmann T, Burgmann H, Makristathis A, et al. *A risk prediction model for screening bacteremic patients: a cross sectional study. PLoS One* 2014;**9**(9):e106765. <https://doi.org/10.1371/journal.pone.0106765>. PMID: 25184209; PMCID: PMC4153716.
24. Choi DH, Hong KJ, Park JH, Shin SD, Ro YS, Song KJ, et al. *Prediction of bacteremia at the emergency department during triage and disposition stages using machine learning models. Am J Emerg Med* 2022;**53**:86–93. <https://doi.org/10.1016/j.ajem.2021.12.065>. Epub 2022 Jan 1. PMID: 34998038.

25. Boerman AW, Schinkel M, Meijerink L, van den Ende ES, Pladet LC, Scholtemeijer MG, et al. *Using machine learning to predict blood culture outcomes in the emergency department: a single-centre, retrospective, observational study. BMJ Open* 2022;**12**(1):e053332. <https://doi.org/10.1136/bmjopen-2021-053332>. PMID: 34983764; PMCID: PMC8728456.
26. Schinkel M, Boerman AW, Bennis FC, Minderhoud TC, Lie M, Peters-Sengers H, et al. *Diagnostic stewardship for blood cultures in the emergency department: a multicenter validation and prospective evaluation of a machine learning prediction tool. EBioMedicine* 2022;**82**:104176. <https://doi.org/10.1016/j.ebiom.2022.104176>. Epub 2022 Jul 16. PMID: 35853298; PMCID: PMC9294655.
27. Collins GS, Reitsma JB, Altman DG, Moons KG. *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ* 2015;**350**:g7594.
28. Charlson ME, Pompei P, Ales KL, MacKenzie CR. *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis* 1987;**40**(5):373–83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8). PMID: 3558716.
29. Dregmans E, Kaal AG, Meziyerh S, Kolfschoten NE, van Aken MO, Schippers EF, et al. *Analysis of variation between diagnosis at admission vs discharge and clinical outcomes among adults with possible bacteremia. JAMA Netw Open* 2022;**5**(6):e2218172. <https://doi.org/10.1001/jamanetworkopen.2022.18172>. PMID: 35737389; PMCID: PMC9226997.
30. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. New York: Springer; 2009.
31. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. *Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res* 2011 Mar;**20**(1):40–9. <https://doi.org/10.1002/mpr.329>. PMID: 21499542; PMCID: PMC3074241.
32. Vickers AJ, van Calster B, Steyerberg EW. *A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res* 2019 Oct 4;**3**:18. <https://doi.org/10.1186/s41512-019-0064-7>. PMID: 31592444; PMCID: PMC6777022.
33. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. *Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol* 2018;**74**(6):796–804. <https://doi.org/10.1016/j.eururo.2018.08.038>. PMID: 30241973; PMCID: PMC6261531.
34. Kerr KF, Brown MD, Marsh TL, Janes H. *Assessing the clinical impact of risk models for opting out of treatment. Med Decis Mak* 2019 Feb;**39**(2):86–90. <https://doi.org/10.1177/0272989X18819479>. Epub 2019 Jan 16. PMID: 30649998; PMCID: PMC6374190.
35. Lien F, Lin HS, Wu YT, Chiueh TS. *Bacteremia detection from complete blood count and differential leukocyte count with machine learning: complementary and competitive with C-reactive protein and procalcitonin tests. BMC Infect Dis* 2022 Mar 26;**22**(1):287. <https://doi.org/10.1186/s12879-022-07223-7>. PMID: 35351003; PMCID: PMC8962279.
36. Steyerberg EW, Eijkemans MJ, Harrell Jr FE, Habbema JD. *Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med* 2000 Apr 30;**19**(8):1059–79. [https://doi.org/10.1002/\(sici\)1097-0258\(20000430\)19:8<1059::aid-sim412>3.0.co;2-0](https://doi.org/10.1002/(sici)1097-0258(20000430)19:8<1059::aid-sim412>3.0.co;2-0)
37. Van Calster B, Steyerberg EW, Collins GS. *Artificial intelligence algorithms for medical prediction should be nonproprietary and readily available. JAMA Intern Med* 2019 May 1;**179**(5):731. <https://doi.org/10.1001/jamainternmed.2019.0597>. PMID: 31058938.

38. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. *Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review*. *BMC Med Res Method* 2022 Dec 12;**22**(1):316. <https://doi.org/10.1186/s12874-022-01801-8>. PMID: 36510134; PMCID: PMC9742671.
39. de Hond AAH, Shah VB, Kant IMJ, Van Calster B, Steyerberg EW, Hernandez- Boussard T. *Perspectives on validation of clinical predictive algorithms*. *NPJ Digit Med* 2023 May 6;**6**(1):86. <https://doi.org/10.1038/s41746-023-00832-9>. PMID: 37149704; PMCID: PMC10163568.
40. Buendía JA, Guerrero Patiño D. *Cost-effectiveness of procalcitonin for detection of serious bacterial infections in children presenting with fever without source*. *BMC Pediatr* 2022 Apr 26;**22**(1):226. <https://doi.org/10.1186/s12887-022-03293-3>. PMID: 35473509; PMCID: PMC9040337.
41. Voermans AM, Mewes JC, Broyles MR, Steuten LMG. *Cost-effectiveness analysis of a procalcitonin-guided decision algorithm for antibiotic stewardship using real-world U.S. hospital data*. *OMICS* 2019 Oct;**23**(10):508–15. <https://doi.org/10.1089/omi.2019.0113>. Epub 2019 Sep 11. PMID: 31509068; PMCID: PMC6806362.



Chapter 4

Bacteremia prediction models to reduce unnecessary blood cultures: external validation in a large US emergency department

Anna Kaal, Laraine Washer, Anastasia Wasylyshyn, Ewout Steyerberg, Prashant Mahajan, Cees van Nieuwkoop

ABSTRACT

Bacteremia prediction models can avoid unnecessary blood cultures while missing few bacteremia patients in European cohorts. We aimed to assess the performance of these models in the United States. We externally validated five prediction models in patients ≥ 18 years old evaluated in an academic emergency department (ED) in the United States, between October 2013 and October 2023. Patients needed to have one or more blood cultures and a procalcitonin (PCT) test within 6 hours of obtaining blood cultures. The models used basic clinical information such as age and vital signs, with or without the addition of C-reactive protein (CRP), PCT or both. Missing data were imputed. The outcome of the models was bacteremia, defined as at least one true positive blood culture obtained within 48 hours of the ED visit. Performance of the models was assessed with the C-statistic for discrimination and calibration plots for calibration. Clinical usefulness was evaluated by the potential reduction in blood cultures, missed bacteremic episodes and decision curve analysis.

We included 12,864 patients (55% male, median age 66 years). 1164 out of 12,864 (9.0%) had bacteremia, while 843 (6.6%) patients had contaminated blood cultures. We observed substantial improvement in performance for models that included PCT, compared to those based on clinical information alone or with CRP. The best performing models included PCT with or without clinical information, with C-statistics of 0.80 (95% CI 0.79-0.82 and 0.79-0.81), adequate overall calibration (predicted 10.4%-13.0%, observed 9.0%). Both models were able to reduce ~30% of blood cultures while missing 5% of bacteremic episodes.

We conclude that previously proposed bacteremia prediction models demonstrated robust applicability in a large United States ED population. Models including procalcitonin have promise to safely avoid many blood cultures in the ED. Further validation and impact studies may support the implementation of these models.

INTRODUCTION

According to the Surviving Sepsis Campaign guidelines it is considered best practice to obtain blood cultures in patients with suspected sepsis before starting antimicrobial therapy.¹ Therefore, it is customary at emergency departments (EDs) to maintain a low threshold to perform blood cultures, as the signs and symptoms of sepsis are nonspecific and may mimic a variety of diseases.¹ Only 4-11% of all blood cultures obtained at the ED yield a true positive pathogen, while up to 50% of all positive cultures may be contaminated.²⁻⁷ Contaminated blood cultures are associated with increased length of stay and more exposure to antibiotics. Additionally, blood cultures are costly (around 90-120 USD per set).⁵⁻⁷

Previously, we developed, validated, and updated multiple bacteremia prediction models based on two European cohorts, with the goal to identify ED patients at low risk for bacteremia and hence, safely reduce unnecessary blood cultures.⁸ The optimal model consisted of routinely available clinical information combined with the biomarker procalcitonin (PCT). This model calculates the individual pre-test probability of bacteremia of adult ED patients with clinical suspicion of infection. Guiding cultures with this model at a 5% risk threshold might avoid a quarter of all blood cultures while only missing 1% of all true positive cultures.⁸

Although these results were very promising, we observed geographical variation in performance, and risks in the lower risk-ranges were slightly overestimated. In this study, we aimed to assess the performance of these models in a large emergency department in the United States through a retrospective chart review.

METHODS

Data and participants

We extracted data from the electronic health records of all patients aged 18 years or older who had a blood culture taken during their ED visit of the University of Michigan Hospital, Ann Arbor, United States, from October 2013 until October 2023. In case of multiple ED visits only the first was included. We excluded patients who left the ED before they received treatment, those whose blood culture was taken before the ED visit or if the first culture was taken more than 24 hours after ED arrival. Patients suspected of endocarditis or those with an intravascular catheter were excluded since we would not recommend using the models in these patient categories. The TRIPOD guideline was followed for reporting on the validation of multivariable prediction models.⁹

Outcome

The outcome of the prediction models was bacteremia, defined as at least one true positive blood culture taken within 48 hours of the start of the ED visit. In a multidisciplinary team consisting of multiple Infectious Disease specialists and hospital epidemiologists, we developed a flowchart to categorize blood cultures with growth of any organism(s) in the following categories: positive, inconclusive, negative and contaminated. We performed a chart review for all inconclusive cultures to evaluate whether the treating clinicians had considered the culture to be truly positive. The flowchart and all retrieved micro-organisms are listed in the Supplement. In summary, Coagulase-negative staphylococci (CNS), *Bacillus* spp., *Corynebacterium* spp., *Cutibacterium* spp., and *Lactobacillus* spp., were considered contaminants unless there were multiple blood cultures with the same bacterium in a patient with a non-vascular prosthetic device (e.g. prosthetic joint). Other reported outcomes are 30-day mortality, admission, ICU admission and start of antibiotic treatment.

Laboratory methods

Blood cultures were collected in an aerobic and anaerobic BD BACTEC bottle and incubated for a minimum of five days using the BD Bactec FX Blood Culture System. Since September 2022, a FilmArray Blood Culture Identification PCR Panel (BCID2) was run on the first positive set of bottles. PCT was first tested using the VIDAS BRAHMS PCT assay (Enzyme Linked Fluorescence Assay), range 0.05 to 200 ng/mL, through May 2018. After that moment, the Roche PCT Assay (Electrochemiluminescence Immunoassay) was used, range 0.02 to 100 ng/mL. CRP was tested with a turbidimetric immunoassay.

Predictors

The predictors used in the models were age, systolic blood pressure (mmHg), heart rate (beats/min), respiratory rate (breaths/min), temperature (°Celsius), Glasgow Coma Scale, CRP (mg/L) and PCT (ng/mL). Biomarker values closest to blood culture collection time were used, with a maximum six-hour difference. For vital signs, the average value during the ED visit up to the moment of blood culture collection was used.

Definitions

We defined suspected endocarditis as the registration of an ICD-10 code during the ED visit related to endocarditis. We created lists of ICD-10 codes for comorbidities, categorized into renal disease, diabetes mellitus, haematological disease,

neurological disease, cardiovascular disease and malignancy (See Supplement). Finally, we selected ICD-10 codes related to an indwelling vascular catheter and to prosthetic material (intravascular and joint). Recent use of antibiotics was defined as use in the seven days before ED visit, and use of immunosuppression as a minimum of two dispenses of immunosuppressive medication in the year before ED visit, of which the last dispense was within the previous six months (See Supplement for lists of antibiotics and immunosuppressants).

Statistical analysis

Continuous variables are presented as mean and standard deviation (SD) or median and interquartile range (IQR). Categorical data are summarized as counts and percentages. Comparisons were made with the χ^2 test, the unpaired t-test or Mann-Whitney U test, as appropriate. We collected data available from the last 10 years, aiming for as many events as possible, with a minimum of 200 events for reliable validation.¹⁰ We performed single imputation of missing data for all predictors except for PCT using multiple imputations by chained equations. From this dataset, we selected all patients with a known PCT value. We validated the European models on the US data, by calculating predictions using the 'predict' function in the R package 'stats' (eTable 1). Predictive performance was assessed by discrimination and calibration. Discriminative ability was assessed by the C-statistic, which is equivalent to the area under the ROC curve.¹¹ To assess whether differences in C-statistic were due to a difference in case-mix or due to misfit of the original model we calculated the model-based C-statistic, which provides the C-statistic for the case-mix in the validation data if the model were fully valid.¹² Calibration was evaluated graphically by calibration plots, with intercept and slope as summary statistics (ideal values 0 and 1 respectively).

Additionally, we explored updating strategies by splitting the database 70%/30%. The 70% part was used to develop multiple updated versions: the original model with updated intercept, intercept and slope and a refitted model with equal weights per dataset (Netherlands, Spain and USA). In this refitted model, an interaction term for PCT and database was added if it improved the model.¹³ All models were then evaluated in the 30% dataset. Other analyses were subgroup analyses for patients using immunosuppressants, for gender, and a complete case analysis. P-values < 0.05 were considered statistically significant. Analyses were performed using RStudio version 2024.09.0.

Potential clinical usefulness

To express clinical usefulness, we calculated the number of avoided blood cultures and number of missed positive cultures at multiple risk thresholds. We report outcomes at a risk threshold of 2.5% and 5% for comparability with the original study.⁸ This threshold reflects the rate of missed bacteremia that is considered acceptable. We consider a plausible range of risk thresholds from 1% to 10%, arguing that any clinician would order blood cultures if the risk would exceed 10%. For PCT, we also reported outcomes at predefined cut-offs of <0.05 ng/mL, <0.10 ng/mL, <0.25 ng/mL and <0.50 ng/mL (Supplement eTable 2 and eTable 3). Clinical usefulness was summarized as Net Benefit (NB), expressed as net interventions avoided per 100 patients, also known as $NB_{\text{Opt-out}}$.¹⁴ This measure quantifies the potential impact on patient outcomes by decision making based on a prediction model. Reference strategies are testing all or none of the patients.

Development versus validation

We assured that inclusion criteria and definitions of predictors and outcome were comparable to the development data. We used histograms to inspect differences in the distribution of continuous predictors with the European data (Supplemental Material). In the Dutch and Spanish setting, blood cultures were ordered based on the judgement of the attending clinician. In Michigan, four screening questions were used to guide whether cultures would be obtained during triage (See Supplement) which could lead to a broader selection of patients. In the Spanish data, patients without a diagnosis of infection within 30 days were excluded. Finally, PCT was tested for all patients in the development cohorts, but only ordered based on the judgement of the attending physician in Michigan.

Ethics

This study was approved by the Institutional Review Board of the University of Michigan (HUM00236866). The requirement for informed consent was waived by the University of Michigan ethics committee due to the retrospective nature of the study. All methods were carried out in accordance with relevant guidelines and regulations.

RESULTS

We identified 46,931 unique ED patients with blood cultures. PCT values were available for 12,864 patients (27%) fulfilling the inclusion criteria. The median age was 65 years, and 7,089/12,864 (55%) were male (Table 1). Bacteremia was diagnosed in 1164 out of 12,864 patients (9.0%), while 843 cultures were

contaminated (6.6%) (Figure 1). The analysis contrasted the 1164 patients with true bacteremia to all others (Negative or contaminated cultures).

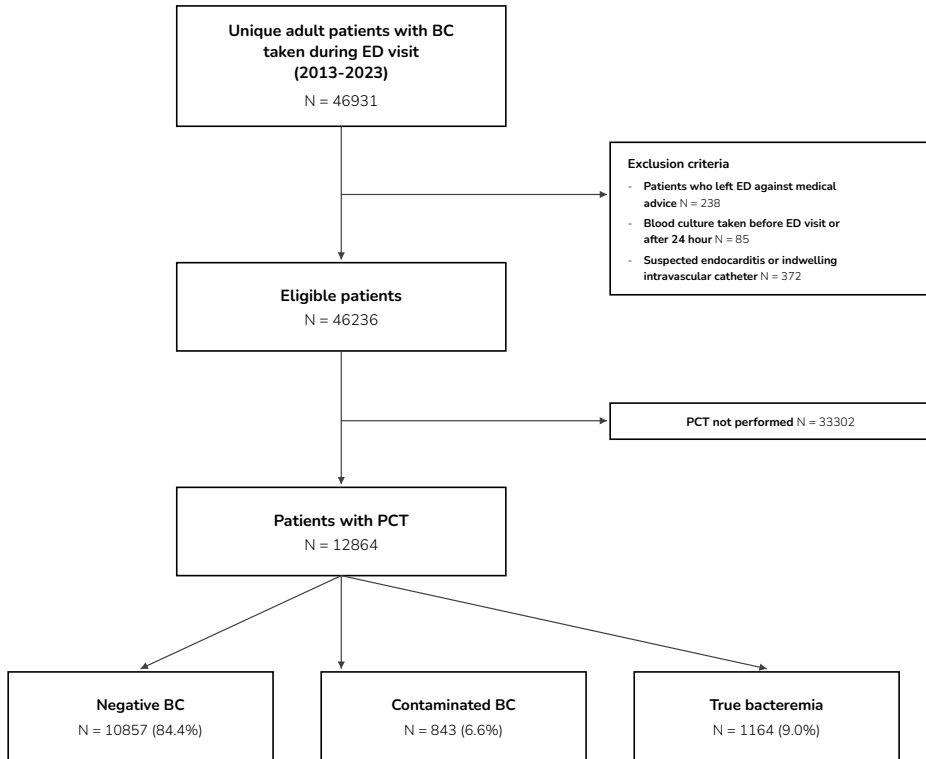


Figure 1. Flowchart of patient inclusions.

The patients in Michigan were comparable to the European patients, and especially similar to the Dutch cohort. (eTable 4) They were less severely ill as reflected in a lower temperature, lower heart rate and lower breathing rate. Patients with a known PCT compared to those without PCT were older, had more comorbidity and suffered from more severe disease as reflected in their vital functions and clinical outcomes. (eTable 5)

Table 1. Baseline characteristics

Characteristics	No bacteremia n = 11700	Bacteremia n = 1164	Total n = 12864	p
Demographics and medical history				
Age, years, (median [IQR])	64.6 [51.5, 75.1]	65.5 [54.4, 75.2]	64.8 [51.7, 75.1]	0.030
Sex (Male)	6374 (54.5)	715 (61.4)	7089 (55.1)	<0.001
Sex (Female)	5326 (45.5)	449 (38.6)	5775 (44.9)	<0.001
Nursing home resident	1208 (10.3)	168 (14.4)	1376 (10.7)	<0.001
Cardiovascular disease	5266 (45.0)	588 (50.5)	5854 (45.5)	<0.001
Diabetes mellitus	3610 (30.9)	409 (35.1)	4019 (31.2)	0.003
Hematological disease	1728 (14.8)	174 (14.9)	1902 (14.8)	0.904
Malignancy	4733 (40.5)	512 (44.0)	5245 (40.8)	0.021
Neurological disease	1920 (16.4)	195 (16.8)	2115 (16.4)	0.796
Pulmonary disease	4929 (42.1)	459 (39.4)	5388 (41.9)	0.081
Renal disease	2799 (23.9)	327 (28.1)	3126 (24.3)	0.002
No comorbidities	1500 (12.8)	122 (10.5)	1622 (12.6)	0.025
Use of immunosuppressants	2213 (18.9)	243 (20.9)	2456 (19.1)	0.113
Recent use of antibiotics	1422 (12.2)	130 (11.2)	1552 (12.1)	0.349
Vital functions in ED				
Breathing frequency, breaths/min, (median [IQR])	20.0 [18.0, 23.0]	20.0 [18.0, 23.2]	20.0 [18.0, 23.0]	0.573
Glasgow coma scale, (median [IQR])	15.0 [14.5, 15.0]	15.0 [14.0, 15.0]	15.0 [14.5, 15.0]	<0.001
Heart frequency, beats/min, (mean (SD))	97.8 (21.5)	104.7 (22.9)	98.4 (21.7)	<0.001
Saturation, %, (median [IQR])	96.0 [93.5, 97.8]	96.0 [93.9, 97.8]	96.0 [93.5, 97.8]	0.171
Systolic blood pressure, mmHg, (mean (SD))	127.3 (26.6)	118.1 (27.3)	126.5 (26.8)	<0.001
Temperature, °C, (median [IQR])	36.9 [36.6, 37.5]	37.3 [36.8, 38.1]	36.9 [36.7, 37.6]	<0.001
Laboratory values in ED				
Creatinine, µmol/L, (median [IQR])	86.7 [65.4, 128.2]	107.0 [75.2, 170.7]	88.4 [66.3, 131.8]	<0.001
Urea, mmol/L, (median [IQR])	7.1 [5.0, 11.4]	9.6 [6.4, 15.4]	7.5 [5.0, 11.8]	<0.001
Procalcitonin, ng/mL, (median [IQR])	0.2 [0.1, 0.7]	2.5 [0.5, 13.4]	0.2 [0.1, 0.9]	<0.001
C-reactive protein, mg/L, (median [IQR])	68.0 [25.0, 145.0]	145.0 [63.5, 223.0]	74.0 [27.0, 155.0]	<0.001
White blood cell count, x10 ⁹ /L, (median [IQR])	10.3 [6.9, 14.8]	11.8 [7.1, 17.4]	10.4 [6.9, 15.0]	<0.001
Neutrophil count, x10 ⁹ /L, (median [IQR])	7.8 [4.8, 11.9]	10.3 [6.1, 15.1]	8.0 [4.9, 12.2]	<0.001
White blood cell count < 0.5 x 10 ⁹ /L	94 (0.8)	46 (4.0)	140 (1.1)	<0.001

Characteristics	No bacteremia	Bacteremia	Total	p
	n = 11700	n = 1164	n = 12864	
Neutrophil count $<0.5 \times 10^9/L$	184 (1.6)	32 (2.9)	216 (1.7)	0.003
Platelet count, $\times 10^9/L$, (median [IQR])	220.0 [158.0, 298.0]	183.0 [118.0, 257.0]	217.0 [155.0, 295.0]	<0.001
Patient outcomes				
Admission	11078 (94.7)	1136 (97.6)	12214 (94.9)	<0.001
ICU admission	2602 (22.2)	433 (37.2)	3035 (23.6)	<0.001
Start of new antibiotics	9703 (82.9)	1097 (94.2)	10800 (84.0)	<0.001
Duration of antibiotic treatment, days, (median [IQR])	7.3 [2.5, 13.6]	13.8 [10.9, 13.9]	7.9 [2.8, 13.8]	<0.001
Contaminated BCs	843 (7.2)	0 (0.0)	843 (6.6)	<0.001
30 day mortality	1410 (12.1)	215 (18.5)	1625 (12.6)	<0.001

Table 1. Baseline characteristics for all included patients. Data are presented as number (percentage) of patients unless otherwise indicated. Recent use of antibiotics: in prior 7 days. IQR = Interquartile Range, ED = emergency department, SD = standard deviation, ICU = Intensive Care Unit, BCs = Blood cultures. Recent use of antibiotics: in the previous seven days. Missing values: Systolic blood pressure: 215 (1.7%), Heart frequency: 144 (1.1%), Breathing frequency: 303 (2.4%), Temperature 843 (6.6%), Glasgow coma scale: 3195 (24.8%), Saturation: 275 (2.1%), Creatinine: 52 (0.4%), Urea: 52 (0.4%), C-reactive protein: 10547 (82%), White blood cell count: 62 (0.5%), Neutrophil count: 229 (1.8%), Platelet count: 148 (1.2%), Duration of antibiotic treatment: 1207 (9.4%).

Model performance

While all models showed slightly lower performance compared to the European setting upon external validation, the discriminative ability of models with PCT was confirmed with AUC values above 0.8 for models including PCT (Table 2). The difference was partly attributable to a difference in case-mix, as indicated by the model-based concordance around 0.84 rather than around 0.87 in the European setting. Calibration plots showed a minor overestimation of risk (Overall predicted risk (10.1-13.0%) vs observed (9.0%)). (Figure 2) All PCT-containing models performed better than those without PCT. At a risk threshold of 2.5%, the basic model with PCT avoided blood cultures in 3,608 out of 12,864 patients (28%) while missing only 52 out of 1164 (4.5%) bacteremic episodes. Adding CRP did not lead to better performance. The 'PCT only' model led to comparable clinical outcomes, potentially avoiding blood cultures in 30% of patients while missing 5.2% of bacteremic episodes at a 4% risk threshold. This model showed slightly poorer calibration. (Table 2) Net benefit analysis showed most benefit for the PCT-containing models (Figure 3). Updating the original models did not lead to improved performance (eTable 6, eFigures 2 and 3). We observed stable results in the complete case analysis and in a second imputation set (eTable 7 and 8). Subanalyses for patients using immunosuppressants or according to gender showed comparable results (eTable 9).

Clinical outcomes for low-risk patients

We compared clinical outcomes for patients with an estimated risk of bacteremia lower than 2.5% by the basic model with PCT, to those with a higher risk. All outcomes were better for low-risk patients, only contamination rates were comparable (6.5% vs. 6.6%, $p = 0.93$). Low-risk patients had a hospital admission rate of 90% vs. 97% ($p < 0.001$). ICU admission rates were 14% vs. 27% ($p < 0.001$). New antibiotics were started in 75% vs. 88% ($p < 0.001$), bacteremia was present in 1.4% vs. 12.0% ($p < 0.001$) and 30-day mortality was 5% vs. 16% ($p < 0.001$).

Table 2. Performance of the models in the Michigan data

Model	Discrimination			Outcomes at a 2.5% risk threshold		Outcomes at a 5% risk threshold		Calibration	
	Development data (95% CI)	Michigan (95% CI)	Model based C-statistic	Avoided BCs (Among 12864)	Missed bacteremic episodes (Among 1164)	Avoided BCs (Among 12864)	Missed bacteremic episodes (Among 1164)	Intercept (95% CI)	Slope (95% CI)
Basic with PCT	0.87 (0.86-0.88)	0.80 (0.79-0.82)	0.84	3608 (28.1%)	52 (4.5%)	6893 (53.6%)	166 (14.3%)	-0.21 (-0.27 to -0.14)	0.79 (0.75 to 0.84)
Basic with CRP and PCT	0.87 (0.86-0.88)	0.80 (0.78-0.81)	0.84	3557 (27.7%)	55 (4.7%)	6850 (53.2%)	177 (15.2%)	-0.23 (-0.30 to -0.16)	0.78 (0.73 to 0.82)
PCT only	0.87 (0.85-0.88)	0.80 (0.79-0.81)	0.82	1347 (10.5%)	16 (1.4%)	4906 (38.1%)	93 (8.0%)	-0.52 (-0.59 to -0.45)	0.84 (0.79 to 0.89)
Basic with CRP	0.71 (0.69-0.73)	0.69 (0.67-0.70)	0.67	360 (2.8%)	8 (0.7%)	2343 (18.2%)	64 (5.5%)	-0.13 (-0.19 to -0.07)	1.07 (0.97 to 1.18)
Basic without biomarkers	0.70 (0.68-0.71)	0.64 (0.62-0.65)	0.64	47 (0.4%)	0 (0%)	1431 (11.1%)	58 (5%)	-0.13 (-0.19 to -0.07)	0.94 (0.83 to 1.06)

Table 2. Performance of the models in the Michigan data. The discrimination is reported as the C-statistic and compared to the model-based C-statistic. The difference between the C-statistic in the development data and the model-based C-statistic reflects the difference attributable to a difference in case-mix. Clinical outcomes are given at a 2.5% and 5% risk threshold with the corresponding number of patients in whom blood cultures could have been avoided, and the number of missed bacteremic episodes. Predictions were calculated with the database variable at 0 based on visual inspection of the predictor distributions (eTable 1). CI = confidence interval, BCs = blood cultures, CRP = C-reactive protein, PCT = Procalcitonin.

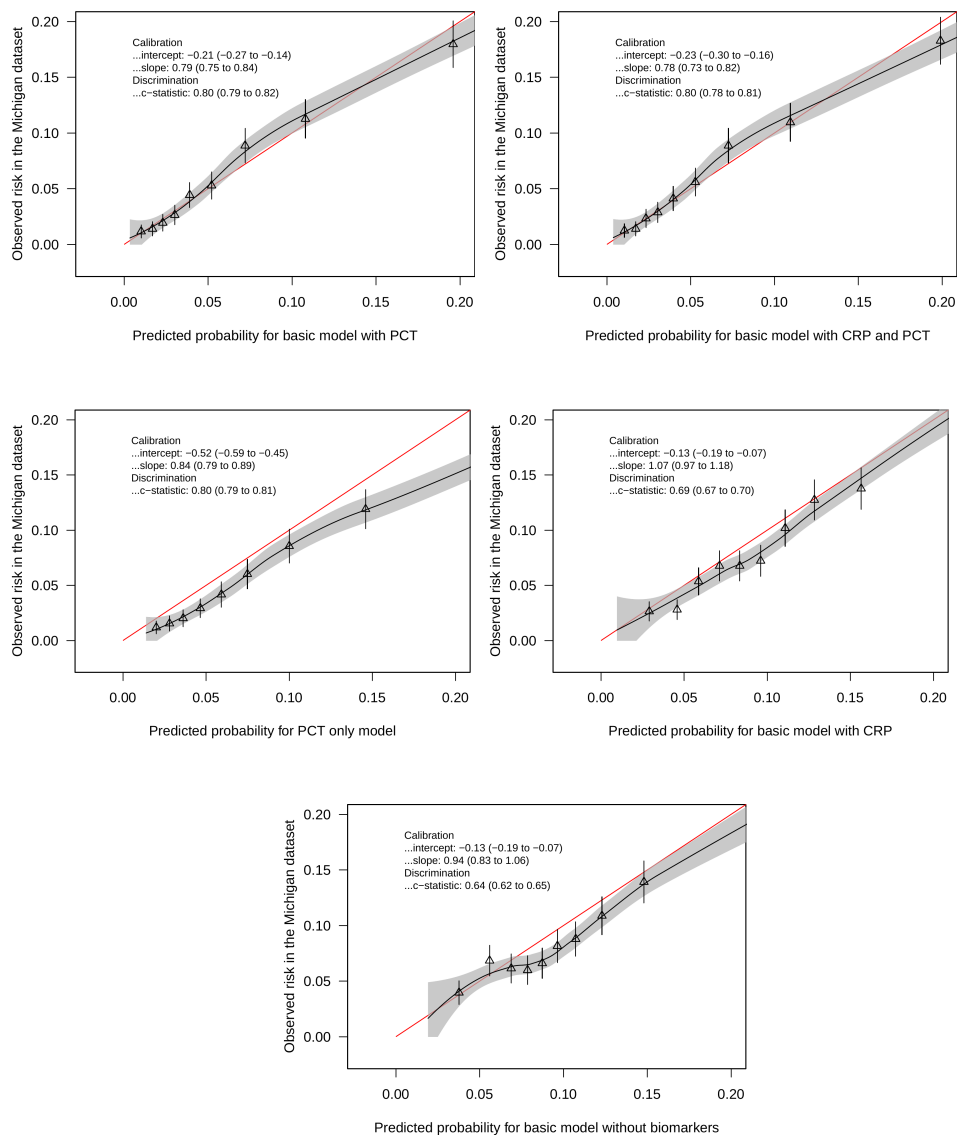


Figure 2: Calibration plots of all models in the Michigan dataset. The X-axis represents the predicted risks, the Y-axis represents the observed risks. CRP = C-reactive protein, PCT = procalcitonin.

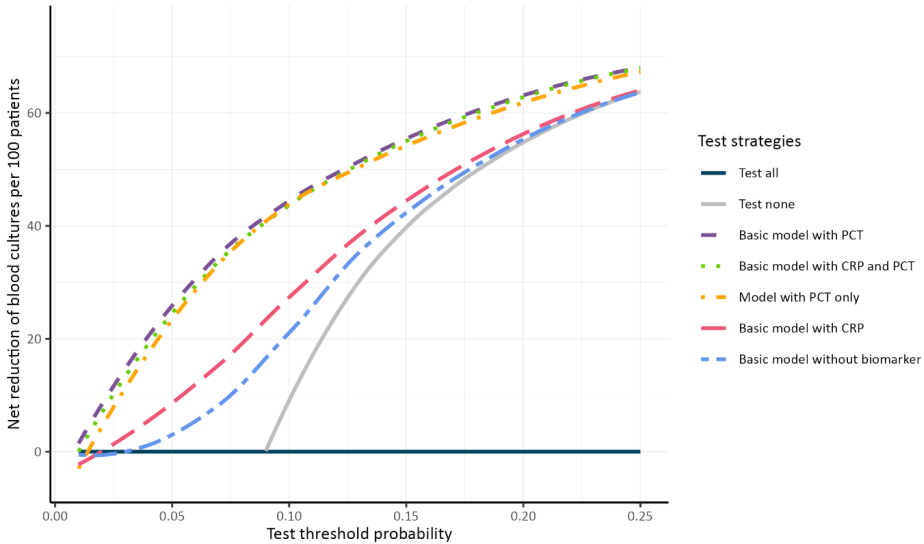


Figure 3: Net benefit curves for all original models. This shows the expected net reduction of blood cultures per 100 patients for all models applied to the Michigan data cohort relative to testing everyone ("Test all") and no one ("Test none"). The x-axis displays the possible risk thresholds for bacteremia on which to decide to withhold blood cultures. For instance: one physician might accept a maximum bacteremia risk of 5%, taking cultures in most patients. The net reduction of blood cultures, or $\text{Net Benefit}_{\text{Opt-out}}$, can be calculated by $\text{TNRR} \times (1-p) - p \times (1-R/R) \times \text{FNRR}$, in which TNR is the true negative rate, R is the risk threshold, p is the disease prevalence and FNR is the false negative rate.^{8,12}

DISCUSSION

We found adequate applicability for diagnostic prediction models for bacteremia in the US. Different model variants were developed and previously validated in Dutch and Spanish patients, and lessons were drawn to create broadly applicable models.⁸ The current validation in US patients confirms the promising performance for models that include PCT to rule out bacteremia. The PCT models were able to avoid over a quarter of blood cultures while missing around 5% of bacteremic episodes, a risk which many consider sufficiently low to withhold blood cultures.^{15,16} It is important to note that the models can also be used with a more cautious approach, according to what the clinician would find acceptable. This would lead to missing fewer bacteremic episodes, and avoiding fewer blood cultures.

To guide blood culture stewardship, indication-based algorithms are considered to achieve the highest impact, with an estimated reduction of 20-50% by avoiding blood cultures in patients with a low risk of bacteremia (< 5-10%).¹⁵⁻¹⁸ Especially within the ED setting, implementation of such algorithms is challenging as diagnoses or source of infection are often still uncertain or even faulty.^{17,19} Some

prediction models aim to predict the risk of bacteremia, something clinicians tend to overestimate.^{20,21} However, many of these tools are too labor-intensive for a busy ED or complex to incorporate into the electronic health record. Moreover, most models lack performance or robust validation, so clinicians are hesitant to use them.²² As overuse of blood cultures is associated with patient harm, robust models that are readily available for application at EDs are urgently needed.

The five models that we developed previously performed slightly worse in the US than in the European context. A part of this difference is explained by case-mix variation. The other part may be due to true differences not captured by these models. For instance, there are differences in clinical settings. In the Netherlands, general practitioners play a major role in treating patients with infectious diseases, and people who are sent to the ED are generally sicker and more often pretreated with antibiotics compared to Spain and the United States. The University of Michigan Hospital is a large center providing academic care, with 1.5 to 2.5 times more patients using immunosuppressants, and few patients with no medical history. However, multiple strategies for local updating did not improve the performance of the models.

The importance of procalcitonin for bacteremia prediction has been shown before in meta-analyses, with a C-statistic of the summary ROC curve of 0.79 for ED patients.²³ Another recent meta-analysis found C-statistics ranging from 0.68 to 0.98.²⁴ For high-quality studies using a PCT cut-off point >0.5 ng/mL, the C-statistic of the summary ROC curve was 0.89. Another large study in the US demonstrated limited sensitivity for 'blood stream infections'. However, using a lower cut-off for PCT, a large number of blood cultures could have been avoided.^{25,26} In our development study, we observed a very strong relation between PCT and bacteremia in the Spanish dataset (C-statistic of 0.9) which motivated an update of the models based on Dutch data only.⁸ In the current validation study, this strong relation was also observed although more in line with the original Dutch development data (C-statistic 0.79).⁸ Surprisingly, basic clinical characteristics did not provide relevant increases in performance, as opposed to the results of the Dutch dataset. Why the predictive performance of PCT differs in different settings is unclear. Although we tried to align all datasets, this could be due to a different patient selection, slight differences in definitions of bacteremia, or timing of the PCT test.

A strength of this study is the large size of the validation dataset with high quality data. We included a broad range of patients, giving a realistic perspective on the application of the models in a US emergency department. A limitation is that we excluded a large part (73%) of eligible patients based on lack of PCT results, without a clear guideline for PCT testing. This has led to a selection of patients

that were more severely ill, as reflected in eTable 5. This selection could have been a confounding factor that could have influenced the results. However, the final selection of patients was more comparable to the development dataset, and the predictive performance of procalcitonin was comparable with the European cohorts in which PCT was tested in all patients. Another limitation is that we used patient chart data, which inherently limits the available information. However, since all predictors used in the models were objective values, we believe this will not have impacted the conclusions of our study. Finally, since the number of positive blood cultures was too high for manual chart review we created a flowchart to categorize blood cultures. In this flowchart, we erred on the side of caution which could have led to an overestimation of the rate of positive cultures. However, if so, this would result in even better clinical outcomes in clinical practice.

Utility in current care

Based on adequate performance and the limited effort required to use a model, we recommend moving forward with either the basic model with PCT or the PCT only model for a prospective validation study. If the setting and case-mix are different from the European and US cohorts, we would advise local validation with model updating if needed. The performance of the models without PCT will not be acceptable for most clinicians. In the Dutch dataset, we observed some improvement in discrimination when adding basic clinical predictors. Ideally, one would evaluate both models using a shadow study, meaning that both models would run in the background without changing diagnostic decisions, and evaluating the performance before moving forward with implementation.¹³ This could also be used to decide whether there is enough added performance to justify the extra effort of implementing the basic model with PCT rather than the PCT only model.

There will be multiple challenges for implementing these models.¹³ We suggest that when the treating clinician considers whether to obtain blood cultures, one would adjust the process as follows. First, if deemed safe to withhold antibiotics for one to three hours one could draw lab testing including PCT and use the model to estimate the risk of bacteremia, using all information available up until the moment that a blood culture would be ordered.²⁷ We propose withholding blood cultures when the estimated risk is <5%, and obtaining them when the risk exceeds 5%. An alternative is to withhold blood cultures if the estimated risk is below the local contamination rate. The model can be adjusted for the clinician's level of risk tolerance, with the consequence that more (or fewer) blood cultures will be withheld. If antibiotics are already needed before model results are available, one could obtain the cultures and decide whether to send cultures for testing based on the model results. In this way,

it is possible to avoid costs from microbiological analysis and negative consequences of contaminated cultures. However, it must be noted that this will only be possible in hospitals where PCT results are known within the first hours. We would argue that if not all predictors are available, one should use the PCT only model, since the added value of clinical predictors was relatively limited compared to the PCT only model. Of note, the models should not be applied to patients with a vascular catheter or suspected endocarditis.

Finally, the implementation of these models in the US could be hindered by issues related to the severe sepsis and septic shock management bundle (SEP-1) compliance.²⁸ SEP-1 is a quality metric that evaluates the timely collection of blood cultures before the administration of antibiotics in patients with suspected severe sepsis and septic shock. However, it only assesses whether cultures were performed, without considering the clinical usefulness of those cultures. Therefore, while our proposed models could lead to more appropriate culture use, the reduction of cultures could negatively impact SEP-1 compliance. In turn, this could lead to decreased financial reimbursement for hospitals, as SEP-1 compliance is tied to value-based purchasing and Medicare reimbursement rates. Addressing these challenges at the policy level is crucial to align quality improvement efforts with financial incentives, ensuring that hospitals can implement diagnostic stewardship models without jeopardizing their reimbursement.

Whether the implementation of these models is worthwhile will differ per setting. For EDs that have already implemented PCT testing in their routine clinical practice, the costs of implementing the models will be low. In settings where PCT testing is not implemented, the costs of the extra PCT tests will have to be weighed against the potential savings of avoiding blood cultures but potentially more missed bacteremia. Costs of PCT, blood cultures and the associated costs with false-negative cultures differ around the world, necessitating local evaluation.

Further research

Important next steps would be to evaluate the potential impact of model implementation, for instance with a shadow study and by evaluating model-user interaction. Questions to be addressed include whether clinicians accept the advice from the model or will the blood cultures still be ordered? What is the outcome of cultures that are taken against the recommendation of the model? And, importantly, does the model impact clinical outcomes?

CONCLUSIONS

Five previously proposed bacteremia prediction models perform well for the US. Especially models including procalcitonin show promising performance to avoid blood cultures in the emergency department while missing few bacteremia cases. Further impact studies may support the implementation of these models in the emergency department.

Data Availability Statement

De-identified, individual patient data supporting the findings of this study are available from the corresponding author upon reasonable request, available from publication without end date. R Markdown files showing the analysis for one of the models are available in the supplementary material.

Use of large language models

ChatGPT was used to improve readability and language.

Potential conflicts of interest

All authors declare no conflicts of interest.

Financial support

This project was made possible because of personal grants appointed to AK, from the Fullbright Commission the Netherlands, the Leids Universitair Fonds, the Catharine van Tussenbroek Fonds, the Cultuurfonds, the Hendrik Mullerfonds and the department of Internal Medicine and Rheumatology, Haga Teaching Hospital, the Netherlands.

Author contributions

Concept and design: AK, ES, CvN. Acquisition of data: AK, PM. Analysis and interpretation of data: AK, LW, AW, ES, PM, CvN. Drafting of the manuscript: AK, critical revision of the manuscript: AK, LW, AW, ES, PM, CvN. Statistical expertise: AK, ES. Acquisition of funding, AK, CvN

SUPPLEMENTARY MATERIAL



REFERENCES

1. Evans L. et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med.* Nov 2021;47(11):1181-1247. doi:10.1007/s00134-021-06506-y
2. Long B, Koyfman A. Best Clinical Practice: Blood Culture Utility in the Emergency Department. *J Emerg Med.* 2016 Nov;51(5):529-539. doi: 10.1016/j.jemermed.2016.07.003. Epub 2016 Sep 14. PMID: 27639424.
3. Coburn B, Morris AM, Tomlinson G, Detsky AS. Does this adult patient with suspected bacteremia require blood cultures? *JAMA.* 2012 Aug 1;308(5):502-11. doi: 10.1001/jama.2012.8262. Erratum in: *JAMA.* 2013 Jan 23;309(4):343. PMID: 22851117.
4. Nannan Panday RS. et al. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. *PLoS One.* 2019 Mar 21;14(3):e0214052. doi: 10.1371/journal.pone.0214052. PMID: 30897186; PMCID: PMC6428292.
5. Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: importance, implications, interpretation and prevention. *Clin Microbiol Infect.* 2018 Sep;24(9):964-969. doi: 10.1016/j.cmi.2018.03.030. Epub 2018 Apr 3. PMID: 29621616.
6. Dempsey C, Skoglund E, Muldrew KL, Garey KW. Economic health care costs of blood culture contamination: A systematic review. *Am J Infect Control.* 2019 Aug;47(8):963-967. doi: 10.1016/j.ajic.2018.12.020. Epub 2019 Feb 20. PMID: 30795840.
7. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization. The true consequences of false-positive results. *JAMA.* 1991 Jan 16;265(3):365-9. PMID: 1984535.
8. Kaal AG. et al. Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model. *J Infect.* Oct 2024;89(4):106251. doi:10.1016/j.jinf.2024.106251
9. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015 Jan 6;162(1):55-63. doi: 10.7326/M14-0697. Erratum in: *Ann Intern Med.* 2015 Apr 21;162(8):600. doi: 10.7326/L15-0078-4. PMID: 25560714.
10. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol.* 2005 May;58(5):475-83. doi: 10.1016/j.jclinepi.2004.06.017. PMID: 15845334.
11. Steyerberg EW. et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010 Jan;21(1):128-38. doi: 10.1097/EDE.0b013e3181c30fb2. PMID: 20010215; PMCID: PMC3575184.
12. Steyerberg EW. (2009) Clinical prediction models: a practical approach to development, validation, and updating. Springer, New York
13. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol.* 2022 Dec 12;22(1):316. doi: 10.1186/s12874-022-01801-8. PMID: 36510134; PMCID: PMC9742671.
14. Kerr KF, Brown MD, Marsh TL, Janes H. Assessing the Clinical Impact of Risk Models for Opting Out of Treatment. *Med Decis Making.* 2019 Feb;39(2):86-90. doi: 10.1177/0272989X18819479. Epub 2019 Jan 16. PMID: 30649998; PMCID: PMC6374190.

15. Ryder JH, Van Schooneveld TC, Diekema DJ, Fabre V. Every Crisis Is an Opportunity: Advancing Blood Culture Stewardship During a Blood Culture Bottle Shortage. *Open Forum Infect Dis*. Sep 2024;11(9):ofae479. doi:10.1093/ofid/ofae479
16. Suleyman G. et al. Blood culture bottle inventory management and clinical conservation during supply shortages. American Society for Microbiology; 2024. https://asm.org/getmedia/8daa65e4-c141-4a0b-9d60-0ae2337505fc/Blood-Culture-shortage-FAQ_FINAL.pdf?ext=.pdf Accessed May 16th, 2025.
17. Fabre V, Sharara SL, Salinas AB, Carroll KC, Desai S, Cosgrove SE. Does This Patient Need Blood Cultures? A Scoping Review of Indications for Blood Cultures in Adult Nonneutropenic Inpatients. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. Aug 22 2020;71(5):1339-1347. doi:10.1093/cid/ciaa039
18. Schoffelen T. et al. European society of clinical microbiology and infectious diseases guidelines for antimicrobial stewardship in emergency departments (endorsed by European association of hospital pharmacists). *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. Jul 17 2024;doi:10.1016/j.cmi.2024.05.014
19. Dregmans E. et al. Analysis of Variation Between Diagnosis at Admission vs Discharge and Clinical Outcomes Among Adults With Possible Bacteremia. *JAMA Netw Open*. 2022 Jun 1;5(6):e2218172. doi: 10.1001/jamanetworkopen.2022.18172. PMID: 35737389; PMCID: PMC9226997.
20. Morgan DJ. et al. Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing. *JAMA Intern Med*. Jun 1 2021;181(6):747-755. doi:10.1001/jamainternmed.2021.0269
21. Schinkel M. et al. Diagnostic stewardship for blood cultures in the emergency department: A multicenter validation and prospective evaluation of a machine learning prediction tool. *EBioMedicine*. 2022 Aug;82:104176. doi: 10.1016/j.ebiom.2022.104176. Epub 2022 Jul 16. PMID: 35853298; PMCID: PMC9294655.
22. Eliakim-Raz N, Bates DW, Leibovici L. Predicting bacteraemia in validated models- a systematic review. *Clin Microbiol Infect*. 2015 Apr;21(4):295-301. doi: 10.1016/j.cmi.2015.01.023. Epub 2015 Feb 10. PMID: 25677625.
23. Hoeboer SH, van der Geest PJ, Nieboer D, Groeneveld AB. The diagnostic accuracy of procalcitonin for bacteraemia: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2015 May;21(5):474-81. doi: 10.1016/j.cmi.2014.12.026. Epub 2015 Jan 14. PMID: 25726038.
24. Julián-Jiménez A, Eduardo García D, Merinos-Sánchez G, García de Guadiana-Romualdo L, González Del Castillo J. Precisión diagnóstica de la procalcitonina para la bacteriemia en el servicio de urgencias: una revisión sistemática [Diagnostic accuracy of procalcitonin for bacteremia in the emergency department: a systematic review]. *Rev Esp Quimioter*. 2024 Feb;37(1):29-42. Spanish. doi: 10.37201/req/099.2023. Epub 2023 Dec 7. PMID: 38058128; PMCID: PMC10874657.
25. Lawandi A et al. Reliability of Admission Procalcitonin Testing for Capturing Bacteremia Across the Sepsis Spectrum: Real-World Utilization and Performance Characteristics, 65 U.S. Hospitals, 2008-2017. *Crit Care Med*. 2023 Nov 1;51(11):1527-1537. doi: 10.1097/CCM.0000000000005968. Epub 2023 Jul 3. PMID: 37395622.

26. Schuetz P. The Value of a Biomarker Should be Judged on What it Adds to the Clinical Assessment: Not the Area Under the Curve. *Crit Care Med*. 2024 Jan 1;52(1):e30-e31. doi: 10.1097/CCM.0000000000006035. Epub 2023 Dec 14. PMID: 38095530.
27. Venkatesh B. et al. Impact of 1-hour and 3-hour sepsis time bundles on patient outcomes and antimicrobial use: A before and after cohort study. *Lancet Reg Health West Pac*. 2021 Nov 2;18:100305. doi: 10.1016/j.lanwpc.2021.100305. PMID: 35024649; PMCID: PMC8654968.
28. Quality Measures Fact Sheet. Severe Sepsis and Septic Shock: Management Bundle Measure (NQF #0500). CMS, August 2020. <https://www.cms.gov/priorities/innovation/media/document/bpci-advanced-alt-fs-my4-sepsis>. Accessed on December 1st, 2025.





Part II

Risk stratification in patients with infectious diseases at the emergency department



Chapter 5

Analysis of variation between diagnosis at admission vs discharge and clinical outcomes among adults with possible bacteremia

Emma Dregmans*, Anna G. Kaal*, Soufian Meziyerh, Nikki E. Kolfschoten, Maarten O. van Aken, Emile F. Schippers, Ewout W. Steyerberg, Cees van Nieuwkoop

*Both authors contributed equally

KEY POINTS

Question

What is the incidence of incorrect identification of site of infection at time of admission, and is this misidentification associated with clinical outcomes?

Findings

In this cohort study of 1477 patients admitted for suspected bacterial infection, misdiagnosed site of infection occurred in 12% of hospitalized patients. No association was found between misdiagnosed site of infection and 30-day mortality, intensive care unit admission, and length of stay, which might be explained by the use of broad-spectrum antibiotics.

Meaning

These findings suggest that misdiagnosed site of infection occurs in every 1 of 9 hospitalized patients, but it is not associated with worse clinical outcomes.

ABSTRACT

Importance

Misdiagnosis of infection is among the most commonly made diagnostic errors and is associated with increased morbidity and mortality. Little is known about how often misdiagnosed site of infection occurs and its association with clinical outcomes.

Objectives

To evaluate the discrepancy between admission and discharge site of infection diagnoses among patients with suspected bacteremia, to explore factors associated with discrepant diagnoses, and to evaluate the association with clinical outcomes.

Design, setting and participants

This cohort study used electronic records of 1477 adult patients who were admitted to the hospital for suspected bacteremia from April 1, 2019, to May 31, 2020, and who had blood cultures taken at the emergency department at Haga Teaching Hospital, The Hague, the Netherlands. Suspected infection sites were classified into 8 categories at admission and discharge. Misdiagnosed site was defined as a discrepancy between the suspected site of infection at admission and at discharge.

Main outcomes and measures

Clinical outcomes were 30-day mortality, intensive care unit admission, length of hospital stay, and antibiotic use, analyzed with logistic and linear regression. Risk factors for misdiagnosed site were determined using regression analysis.

Results

A total of 1477 patients (820 [55.5%] male; median [IQR] age, 68 [56-78] years) were analyzed. The rate of misdiagnosed site of infection was 11.6% (171 of 1477); 3.1% of all patients (46 of 1477) ultimately had no infection. No association was found between misdiagnosis and 30-day mortality (adjusted odds ratio [aOR], 0.8; 95% CI, 0.3-1.9; $P = .60$), intensive care unit admission (aOR, 1.3; 95% CI, 0.6-3.0; $P = .54$), and hospital length of stay (adjusted increase of stay, 15.5%; 95% CI, -3.1% to 37.7%; $P = .11$). Misdiagnosed site was associated with receiving broad-spectrum antibiotics (aOR, 4.0; 95% CI, 1.8-8.8; $P < .001$). Older age, dementia, a positive urine sediment test result without urinary symptoms, and suspicion of an intravascular, central nervous system, or bone and joint infection were risk factors for misdiagnosed site of infection.

Conclusions and relevance

In this cohort study, misdiagnosed site of infection occurred in 1 of 9 patients and was not associated with worse short-term clinical outcomes. Clinicians should be aware of risk factors associated with misdiagnosed site of infection and potential inappropriate antibiotic use.

INTRODUCTION

Increasing attention is being paid to the prevention of diagnostic errors, including misdiagnoses, because these errors have been associated with inappropriate treatment, harm to patients, and increasing costs.¹⁻³ Little is known about the incidence of misdiagnoses in hospitals and their consequences.⁴ The common overall estimated percentage of diagnostic error is 10% to 15%, but great variation exists among studies because of a lack of unified terminology and different settings of studies.⁴⁻⁹ In this respect, emergency departments (EDs) are known to have higher rates of misdiagnoses because of the complex decision-making process, limited available clinical information, high workload, and time constraints.^{2,9-11}

A study¹² among tertiary-level EDs in Japan on the misdiagnosis of infections found that 11.6% (113 of 974) of ED patients had a misdiagnosis, with urinary tract infection being the most common misdiagnosis. Misdiagnosed infections may lead to morbidity and mortality, as well as to inappropriate diagnostic testing and antimicrobial treatment, and thus contribute to increasing antimicrobial resistance.^{4,11,13,14} Nonspecific symptoms related to infections make it challenging to establish the correct diagnosis. Fever, for example, is a symptom of low specificity and may indicate a self-limiting infection as well as severe sepsis.^{15,16}

Identifying risk factors for misdiagnosis might be helpful to improve diagnostic accuracy. Previous studies^{9,10} have primarily focused on assessing the type of error in the diagnostic process and on the predictive role of context (eg, noise) and work experience in misdiagnoses. Data on patient factors associated with misdiagnoses are lacking.

We recently performed an observational study¹⁷ of all adults presenting to the ED in a large teaching hospital in the Netherlands from whom a blood culture was taken because of suspected bacteremia. The aims of the current substudy were to evaluate the rate of misdiagnosed site of infection in ED patients with suspected bacteremia, to examine the association of misdiagnosis with clinical outcomes, and to identify risk factors associated with misdiagnosis.

METHODS

Study Design and Setting

The current study is a substudy of PredictED (Predicting Microbiological and Clinical Outcomes of Patients With Suspected Bacteremia Presenting at Emergency Department: A Prospective Observational Study),¹⁷ a single-center observational cohort study that was conducted at the ED of the Haga Teaching Hospital in

The Hague, the Netherlands, a large hospital with approximately 600 beds and 50000 ED visits per year. A total of 2333 adult patients who presented at the ED of HagaTeaching Hospital and had a blood culture taken from April 1, 2019, to May 31, 2020, were included in the study. For this substudy, patients were excluded if they were not hospitalized or transferred to another hospital during their admission, because in such cases the final diagnosis was not available. In addition, patients who were not diagnosed with any infection by the ED physician or patients with an unidentified site of infection at the ED were excluded. PredictED was approved by the Haga Teaching Hospital Medical Ethics Committee and the Haga Teaching Hospital Institutional Scientific Review Board. The need for informed consent was waived because this is an observational quality control study without any effect on patients' medical management. The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Data Collection

Demographic characteristics and clinical and biochemical findings were collected from the patient electronic records by 2 investigators (E.D. and A.G.K.). Site of infection was obtained by reviewing the patient record, including the hospital discharge letter. A Cohen κ score was used to assess interrater agreement between the investigators (E.D. and A.G.K.). In case of disagreement, a third internal medicine physician was consulted (S.M.). The primary outcome measures were the rate of misdiagnosed infection, 30-day mortality, and intensive care unit (ICU) admission. Secondary outcomes included the length of hospital stay and the administration of broad-spectrum antibiotics.

Data Definitions

Suspected infection sites were classified into 8 categories: respiratory, urogenital, skin and soft tissue, central nervous system (CNS), abdominal, intravascular, bone and joint, and other infection. The diagnosis at discharge was classified into identical categories as well as unidentified site of infection and no infection. Misdiagnosed site was defined as a discrepancy between the suspected site of infection at admission and at discharge. In case of a possible second infection, a misdiagnosis was defined as a discrepancy between the infection site listed as the reason for hospitalization at admission and at discharge. The quick Sepsis Related Organ Failure Assessment (qSOFA) score was used as a measure for disease severity, and the Charlson Comorbidity Index was used to measure the burden of comorbidity.^{18,19} Urinary symptoms were defined as occurrence of any Loeb criteria (dysuria, suprapubic pain/tenderness, frequency, or urgency).²⁰ A positive urine sediment test result indicated

a positive nitrite test result or more than 5 white blood cells per high-power field. Broad-spectrum antibiotics were defined as intravenous second- or third-generation cephalosporin with or without an aminoglycoside.²¹ Bacteremia was present when a blood culture, taken within 48 hours after presentation at the ED, showed bacterial growth and was not deemed a contaminant. Coagulase-negative staphylococci, *Bacillus* species, *Corynebacterium* species, *Cutibacterium* species, *Viridans* group streptococci, and *Lactobacillus* species were considered contaminants unless associated with an intravascular catheter or device, suspected endocarditis, or the presence of multiple positive blood culture results with the same bacterium.

Statistical Analysis

Descriptive statistics were presented as numbers and percentages for categorical variables. All continuous variables were expressed in medians and IQRs. The χ^2 test or Fisher exact test for categorical variables was used. Continuous variables were tested with the Mann-Whitney U test. A 2-sided $P < .05$ was considered statistically significant. Rates of misdiagnosis were calculated for each site of infection.

Univariate and multivariable regression analyses were performed to investigate the association of misdiagnoses with outcomes. The binary dependent variables were 30-day mortality, ICU admission, and administration of broad-spectrum antibiotics, which were analyzed with logistic regression models. A linear model was used for the length of hospital stay with a logarithmic transformation because of skewed data. Multivariable analyses were performed to correct for age, Charlson Comorbidity Index, qSOFA score, and site of infection at discharge. A sensitivity analysis investigated the association of misdiagnosis with clinical outcomes when patients with no infectious disease at discharge were excluded.

To identify variables associated with misdiagnosed sites, a multivariable logistic regression model was built using the LASSO (least absolute shrinkage and selection operator) method.²² Eighteen candidate predictor parameters were considered for inclusion based on a sample size of 1900 participants, an outcome proportion of 15%, and an expected shrinkage of 10%. These variables were age, sex, dementia, use of immunosuppressants, recent use of antibiotics, altered mental status, systolic blood pressure, breathing frequency, C-reactive protein level, presence of Loeb criteria, positive urine sediment test result without Loeb criteria, and suspected site of infection at the ED (8 site categories).²⁰ Missing data (1.4%) were assumed to occur at random and imputed 10 times using chained equations.²³ The variables that contained missing values were altered mental status, systolic blood pressure, breathing frequency, C-reactive protein level, and positive urine sediment test result without Loeb criteria. The 10 resulting imputed data sets were stacked into a single

large data set and then incorporated into 1 model through weights. Because of the many different diagnoses within the group labeled as “other infection,” this category was excluded in the LASSO analysis. Internal validation was conducted with a 200-bootstrap resampling procedure, where the LASSO procedure was repeated in each bootstrap sample to estimate the optimism corrected performance.²⁴ Associations were presented as regression coefficients for linear regression and odds ratios (ORs) for binary outcomes. With the use of the LASSO method, model selection through shrinkage is achieved to reduce overfitting. Because of shrinkage, CIs cannot be determined for the final set of ORs, in contrast to the full model, including all candidate variables. Discrimination of the model was assessed with the concordance statistic (C statistic), an equivalent to the area under the receiver operating characteristic curve. Statistical analyses were performed using R software, version 1.3.1073 (R Foundation for Statistical Computing) and IBM SPSS Statistics, version 26 (IBM Inc).

RESULTS

Between April 2019 and May 2020, a total of 2333 patients presented at the ED of Haga Teaching Hospital and had a blood culture taken. Of these, 1477 patients were admitted to the hospital with a suspected infection (eFigure in the Supplement). The median age was 68 years (IQR, 56-78 years), 657 (44.5%) were female, and 820 (55.5%) were male. A total of 171 admitted patients (11.6%) had a misdiagnosed site of infection (Table 1). The Cohen κ for interrater agreement of site of infection was 0.87.

Patient Characteristics

Patients with misdiagnosed site of infection were older (median age, 72 [IQR, 64-80] vs 68 [IQR, 55-78] years; $P < .001$) (Table 1); more likely to have a history of cardiovascular disease (96 [56.1%] vs 544 [41.7%]; $P < .001$), kidney disease (34 [19.9%] vs 151 [11.6%]; $P = .002$), or dementia (14 [8.2%] vs 50 [3.8%]; $P = .008$); and had a higher Charlson Comorbidity Index (median, 4 [IQR, 1-6] vs 3 [IQR, 2-5]; $P = .003$). Initial vital signs were similar across groups. More patients with misdiagnosed site of infection had altered mental status (28 [16.8%] vs 105 [8.9%]; $P = .001$), a positive urine sediment test result (74 [43.3%] vs 429 [32.8%]; $P = .007$), and bacteremia (42 [24.6%] vs 218 [16.7%]).

Table 1 Characteristics of patients with misdiagnosed and correctly diagnosed site of infection (n= 1477)^a

Characteristics ^{b-e}	Misdiagnosed 171	Correctly diagnosed 1306	p-value
Demographics and medical history			
Age (years)	72 [64-80]	68 [55-78]	<0.001
Male, n (%)	89 (52)	731 (56)	0.33
Female, n (%)	82 (48)	575 (44)	
Nursing home resident, n (%)	18 (10.5)	143 (10.9)	0.87
History of cardiovascular disease, n (%)	96 (56.1)	544 (41.7)	<0.001
History of pulmonary disease, n (%)	34 (19.9)	392 (30.0)	<0.001
History of renal disease, n (%)	34 (19.9)	151 (11.6)	0.002
Dementia, n (%)	14 (8.2)	50 (3.8)	0.008
Charlson comorbidity index	4 [2-6]	3 [1-5]	0.003
Vital signs			
Initial SBP (mmHg)	128 [110-150]	129 [112-148]	0.49
Initial HR (/min)	98 [84-115]	102 [89-115]	0.16
Initial BF (/min)	20 [16-25]	20 [16-26]	0.08
Altered mental status, n (%)	28 (16.8)	105 (8.9)	0.001
Initial Body temperature (°C)	38.5 [37.8-39.3]	38.6 [37.9-39.2]	0.40
qSOFA score, n (%)			0.10
0	67 (43.5)	496 (44.7)	
1	59 (38.3)	475 (42.8)	
2	23 (14.9)	118 (10.6)	
3	5 (3.2)	21 (1.9)	
Laboratory results			
CRP (mg/L)	100 [40-183]	95 [35-197]	0.95
WBC (x10 ⁹ /L)	11.3 [8.6-16.1]	12.1 [8.4-16.8]	0.44
Positive urine sediment, n (%) ^c	74 (43.3)	429 (32.8)	0.007
Bacteremia, n (%) ^e	42 (24.6)	218 (16.7)	0.01

Abbreviations: BF, breathing frequency; CRP, C-reactive protein; HR, heart rate; qSOFA, quick Sequential Organ Failure Assessment; SBP, systolic blood pressure; WBCs, white blood cells. SI conversion factors: To convert CRP to milligrams per liter, multiply by 10; WBCs to $\times 10^9/L$, multiply by 0.001. ^a Data are presented as number (percentage) of patients unless otherwise indicated. Systolic blood pressure data are missing in 7 patients, HR data in 4 patients, BF data in 93 patients, altered mental status data in 131 patients, CRP data in 1 patient, WBC count data in 2 patients, data on leukocytes in urine sediment in 6 patients, and qSOFA data in 213 patients. ^b Positive urine sediment test result indicates nitrite positive or more than 5 WBCs per high-power field. ^c Present when within 48 hours after presentation at the emergency department a blood culture yielded a pathogen typical for the infection site.

Rate of Misdiagnoses

The most common sites of infections as diagnosed at admission were respiratory (657 of 1477 [44.5%]) and urogenital (336 of 1477 [22.7%]). Bone and joint (12 of

38 [31.6%]), CNS (8 of 21 [38.1%]), and intravascular (14 of 30 [46.7%]) infections had the highest rates of misdiagnosed site (Figure). The rate of misdiagnoses per initial sites of infection are outlined in the Figure. Thirteen patients had a possible second infection at admission, 5 at both admission and discharge, and 7 at discharge only. Of all 13 patients with a possible second site of infection, 8 (61.5%) were misdiagnosed. Among all 1477 patients, 46 (3.1%) were ultimately discharged with a diagnosis of no infection (Table 2).

Association of Misdiagnosis With Clinical Outcomes

Of 171 patients with a misdiagnosed site of infection, 18 (10.5%) died within 30 days of presentation (Table 3); in patients with correctly diagnosed sites, the 30-day mortality was 8.5% (111 of 1306). Among patients with misdiagnosed sites, 11 (6.4%) were admitted to the ICU, compared with 89 patients (6.8%) with correctly diagnosed sites. These differences were not statistically significant because a misdiagnosed site of infection was not associated with 30-day mortality (adjusted OR [aOR], 0.8; 95% CI, 0.3-1.9; $P = .60$) or ICU admission (aOR, 1.3; 95% CI, 0.6-3.0; $P = .53$). A sensitivity analysis, excluding patients with no infectious diagnosis at discharge, also showed no association between misdiagnosed site of infection and 30-day mortality (aOR, 0.8; 95% CI, 0.3-1.8; $P = .55$) or ICU admission (aOR, 1.3; 95% CI, 0.6-3.0; $P = .51$). For patients with misdiagnosed sites, the median (IQR) hospital length of stay was 5 days (IQR, 3-9 days), compared with 4 days (IQR, 2-7 days) for patients with correctly diagnosed sites ($P = .007$), a 22.3% (95% CI, 5.6%-41.5%) increase in hospital length of stay; however, when adjusted for confounders, this association disappeared (15.5%; 95% CI, -3.1% to 37.7%; $P = .11$). Most patients (149 of 171 [87.1%] with misdiagnosed sites and 1110 of 1306 [84.9%] with correctly diagnosed sites) received antibiotic therapy. A statistically significantly higher percentage of misdiagnosed patients received broad-spectrum antibiotics compared with patients with a correct diagnosis (86 of 112 [76.8%] vs 532 of 917 [58.0%]; aOR, 4.0; 95% CI, 1.8-8.8; $P < .001$).

Variables Associated With Misdiagnosis

A multivariable logistic regression model was used to assess variables associated with a misdiagnosed site of infection. In Table 4, both the candidate variables and the selected variables are shown with corresponding ORs (regression coefficients can be found in the eTable in the Supplement). Four of 13 variables were selected using the LASSO logistic regression: age, dementia, a positive urine sediment test result without Loeb criteria, and site of infection at ED. Discrimination of the overall model was moderate, with a C statistic of 0.70 (95%CI, 0.68-0.70). The performance remained similar after correcting for optimism ($C = 0.69$; 95%CI, 0.64-0.74).

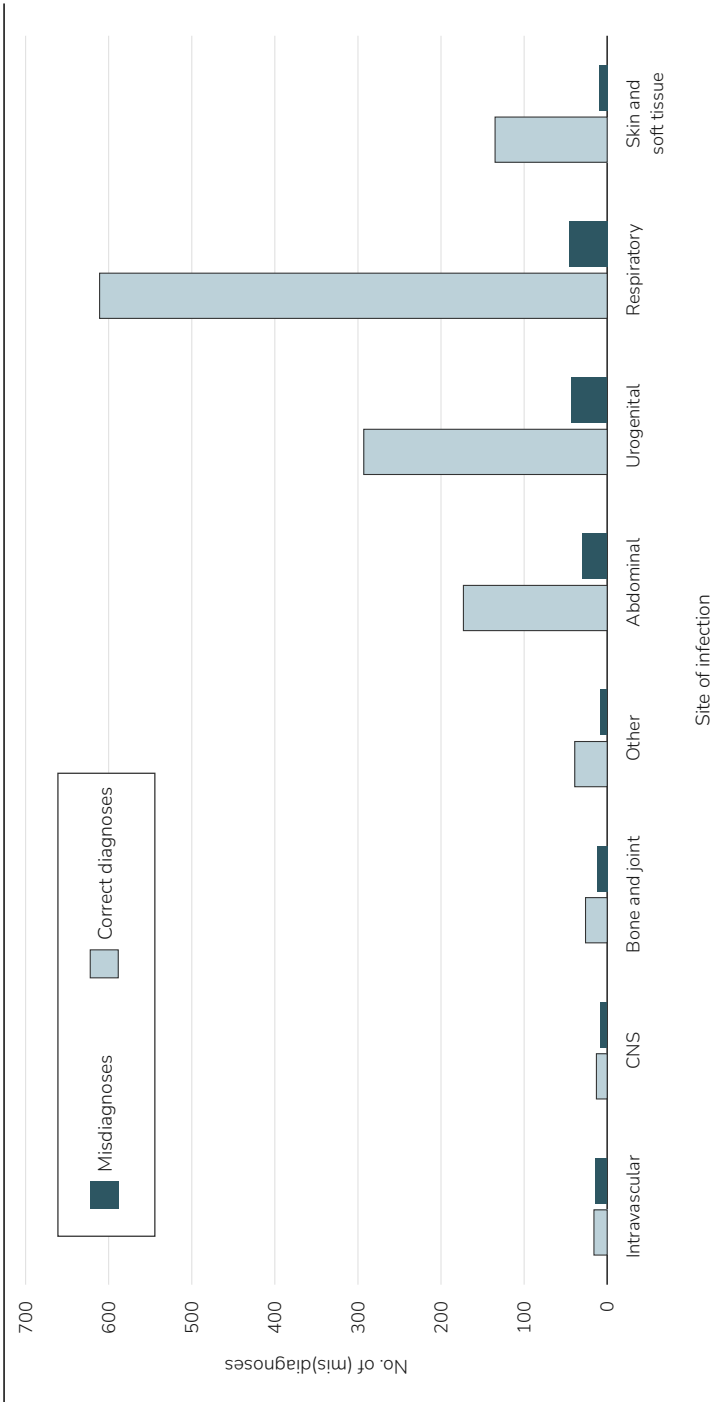


Figure. Misdiagnoses per Site of Infection

Table 2 Rate of misdiagnosis by Site of Infection in the 1477 Study Patients

Site of infection at ED	Misdiagnosed Sites		No. of infections by site of infection									
	No. (%)	(n = 171), No./total	Respiratory	Urogenital	Abdominal	Skin and Soft tissue	Bone and Joint	Intravascular	CNS	Other ^a	No infection	Unidentified
171/1477 (11.6)												
Respiratory	46/657 (7.0)		611	11	6	3	0	0	1	3	15 ^b	7
Urogenital	43/336 (12.8)		9	293	8	3	1	4	2	3	5 ^c	8
Abdominal	30/203 (14.8)		4	10	173	0	0	3	0	1	8 ^d	4
Skin and Soft tissue	10/145 (6.9)		1	1	1	135	4	2	0	0	1 ^e	0
Bone and Joint	12/38 (31.6)		0	1	0	0	26	0	0	1	10 ^f	0
Intravascular	14/30 (46.7)		1	3	2	1	0	16	0	0	4 ^g	3
CNS	8/21 (38.1)		1	2	0	0	1	1	13	2	1 ^h	0
Other ^a	8/47 (17.0)		3	1	0	1	0	0	0	39	2 ⁱ	1

Abbreviations: CNS, central nervous system; ED, emergency department. ^a Other sites of infection (eg, ear, nose, and throat) and viral infection. ^b Pulmonary embolism (n = 2), congestive heart failure (n = 2), carcinoma (n = 2), chemical pneumonia (n = 1), neutropenic fever after chemotherapy (n = 1), fever after radiation therapy (n = 1), absorption fever attributable to hematoma (n = 1), pleural effusion (n = 1), nitrofurantoin-induced acute pulmonary toxicity (n = 1), lytic rib lesion (n = 1), symptoms after thoracic endovascular aortic repair (n = 1), and unexplained syncope (n = 1). ^c Terminal heart failure (n = 1), metastatic malignancy (n = 1), delirium (n = 1), terminal kidney insufficiency (n = 1), and dehydration (n = 1). ^d Hepatic encephalopathy (n = 1), silent jaundice (n = 1), hepatic lymphoma (n = 1), inflammatory bowel disease (n = 1), fever after immunotherapy (n = 1), benign paroxysmal positional vertigo (n = 1), ileus (n = 1), and ischemic colitis (n = 1). ^e Lymphoma (n = 1). ^f Arthrosis (n = 2), muscle complaints (n = 2), tophaceous gout (n = 2), calcium pyrophosphate deposition arthritis (n = 2), psoriatic arthritis (n = 1), and osteoarthritis (n = 1). ^g Disseminated intravascular coagulation in metastatic malignancy (n = 1), portal vein thrombosis (n = 1), mitral valve vegetation without clinical symptoms (n = 1), and suspected endocarditis with negative blood culture. ^h Hepatic encephalopathy (n = 1). ⁱ Febrile neutropenia in acute myeloblastic leukemia (n = 1) and metastatic malignant neoplasm (n = 1).

Table 3 Outcomes of 1477 Patients with Misdiagnosed and Correctly Diagnosed Site of Infection

Outcome	Misdiagnosed site ^b	Correctly diagnosed site ^b	Univariate		Multivariable ^a	
	(n=171)	(n=1306)	OR [95% CI]	p-value	aOR [95% CI]	p-value
30-day mortality, n (%)	18 (10.5)	111 (8.5)	1.3 [0.7-2.1]	0.38	0.8 [0.3-1.9]	0.60
ICU admission, n (%)	11 (6.4)	89 (6.8)	0.9 [0.5-1.8]	0.85	1.3 [0.6-3.0]	0.53
Broad-spectrum antibiotic therapy ^c , n (%)	86/112 (76.8)	532/917 (58.0)	2.4 [1.5-3.8]	<0.001	4.0 [1.8-8.8]	<0.001
Hospital length of stay (days), median (IQR) ^d	5 [3-9]	4 [2-7]	22.3 [5.6-41.5] ^e	0.007	15.5 [-3.1-37.7] ^e	0.11
Sensitivity analysis^f						
30-day mortality, n (%)	10 (8.0)	111 (8.5)	OR [95% CI]	p-value	OR [95% CI]	p-value
			0.9 [0.5-1.8]	0.85	0.8 [0.3-1.8]	0.55
ICU admission, n (%)	9 (7.2)	89 (6.8)	1.1 [0.5-2.2]	0.87	1.3 [0.6-3.0]	0.51

Abbreviations: aOR, adjusted odds ratio; ICU, intensive care unit; OR, odds ratio.

^a Adjusted for age, Charlson Comorbidity Index, quick Sequential Organ Failure Assessment, and site of infection at discharge.^b Data are presented as number (percentage) of patients unless otherwise indicated.^c Intravenous second- or third-generation cephalosporin with or without an aminoglycoside administered within 8 hours after a blood culture was taken.^d Excluding patients who died within 30 days.^e Log-linear model in which the exponentiated regression coefficient represents the percentage of change in geometric means when all other variables are held at a fixed value.^f Excluding patients with no infectious disease at discharge.

DISCUSSION

This cohort study analyzed misdiagnoses of infection site among ED patients hospitalized for suspected infection. We found a misdiagnosis rate of 11.6% but no association between misdiagnosed site of infection and worse clinical outcomes, including 30-day mortality, need for ICU admission, and length of hospitalization. However, more patients with misdiagnosed sites of infection received broad-spectrum antibiotics. Older age, presence of dementia, a positive urine sediment test result without urinary symptoms, and particular suspected sites of infection (intravascular, CNS, and bone and joint) were the most prominent risk factors associated with misdiagnosed site of infection in our study.

Literature on misdiagnosed site of infection among hospitalized ED patients is scarce. To our knowledge, this is the first study to investigate the association between misdiagnoses and clinical outcomes while simultaneously identifying patient risk factors for misdiagnosed site of infection in this specific population. A Japanese study by Abe et al¹² investigated the association of misdiagnosed site of infection with clinical outcomes in 974 patients, using a similar definition. However, Abe et al¹² also included patients with an unidentified site of infection at admission. When these patients were excluded, they found an overall rate of misdiagnoses of 8.1% (76 of 937).

In our study, sites of infection that were less common (eg, intravascular and CNS) were associated with the highest risks for misdiagnosis, which is in line with previous findings.¹² The higher rate of misdiagnoses in more uncommon sites of infections might be explained by a lack of exposure or more difficulty in identifying infections at less common sites. However, we think that the higher rate of misdiagnoses is also a reflection of the more severe consequences of missing a diagnosis in these categories, which leads physicians to err on the side of caution. For instance, it can be challenging to distinguish septic arthritis from a simple gout. Misdiagnosis of abdominal infections was more often reported in our study than in the study by Abe et al¹² (30 of 203 [14.8%] vs 7 of 186 [3.8%]). Another study¹¹ that evaluated misdiagnoses of pneumonia in the ED found no respiratory site of infection in 26 of 195 patients (13.3%) on discharge from the internal medicine ward, in contrast to 46 of 657 (7.0%) misdiagnosed respiratory infections in our study. We hypothesize that besides differences in study populations, the differences are partly explained by the challenges in this field of research to assess records objectively.

Table 4 Full model and final determinants of misdiagnosed site of infection

Variables	Full model OR [95% CI]	Reduced model OR ^{ab}
Age (per decade)	1.2 [1.1-1.3]	1.1
Sex (Male)	0.9 [0.7-1.3]	not selected
Use of immunosuppressants	1.1 [0.6-1.8]	not selected
Use of antibiotics	0.9 [0.6-1.3]	not selected
Altered mental status	1.0 [0.7-1.7]	not selected
SBP	1.0 [1.0-1.0]	not selected
CRP	1.0 [1.0-1.0]	not selected
BF	1.0 [1.0-1.0]	not selected
Occurrence of Loeb criteria	1.3 [0.8-2.2]	not selected
Dementia	1.9 [0.9-3.7]	1.4
Positive urine sediment without Loeb criteria ^{c,d}	1.4 [0.9-2.1]	1.1
Site of infection at ED		
Abdominal	2.2 [1.3-3.7]	1.4
Bone and Joint	6.1 [2.8-13.4]	3.4
CNS	11.0 [4.0-29.7]	5.1
Intravascular	12.0 [5.3-27.2]	6.7
Respiratory	1 [reference]	1 [reference]
Urogenital	1.6 [1.0-2.7]	1.2
Skin and Soft tissue	1.1 [0.5-2.2]	1.0

Abbreviations: BF, breathing frequency; CNS, central nervous system; CRP, C-reactive protein; ED, emergency department; OR, odds ratio; SBP, systolic blood pressure. ^a Model selection is achieved by shrinkage, so the ORs with CIs cannot be determined. ^b The tuning parameter (λ) is chosen by cross-validation.²⁵ The value of λ that gives the simplest model but also lies within 1 SE of the optimal value of λ was selected. ^c The Loeb criteria are dysuria, suprapubic pain or tenderness, frequency, or urgency. A positive urine sediment test result indicates nitrite positive or more than 5 white blood cells per high-power field. ^d Occurred in 36 of 163 misdiagnosed patients (22.1%) and 183 of 1267 nonmisdiagnosed patients (14.4%).

Other studies^{2,10} have shown that diagnostic errors in EDs can result in serious patient harm. A possible explanation for the difference with our study is our focus on patients with infections and suspected bacteremia or sepsis. Most patients with a misdiagnosed site of infection in our study were treated with broad-spectrum antibiotics, following sepsis guidelines. Broad-spectrum antibiotics were likely an effective treatment for various sites of infection, even though an incorrect site was diagnosed. As such, we did not find an association of misdiagnosed infection site with ICU admissions or mortality even though patients with misdiagnosed sites had a higher rate of bacteremia. However, the effects of misdiagnosed site of

infection on clinical outcomes may be much broader and more subtle than what we examined in this study. For example, misdiagnosed site might be associated with additional diagnostic testing. Based on our results, we noticed that 20% of misdiagnosed patients received antibiotics unnecessarily, because the final correct diagnosis was no infection. Unnecessary use of antibiotics may lead to adverse effects for individual patients and may contribute to overall increasing antimicrobial resistance, which is considered a large public health threat.²⁵ Our unadjusted results showed an association between misdiagnosed site of infection and hospital length of stay, but after adjustment for age, severity of disease, and site of infection at discharge, this association disappeared. When adjusting for these confounders one by one, the association disappeared when corrected for site of infection at discharge. In contrast, Abe et al¹² demonstrated that patients with a misclassified site of infection at admission had 2.7-fold higher odds of in-hospital mortality. This difference from our results could be explained by differences in study populations, with their cohort having a higher median age, more patients with severe disease as shown by a qSOFA of 2 or greater (38.7% vs 11.3%), and a higher overall mortality (15% vs 8.7%). More research is needed to clarify this discrepancy in the association of misdiagnosed infections with adverse outcomes. We propose that in future research more microbiological outcomes should be evaluated, preferably including long-term outcomes, such as the emergence of antibiotic resistance in patients receiving inappropriate antibiotics.

In this study, we built a model to help physicians identify patients at high risk of a misdiagnosed site of infection using age, dementia, presence of a positive urine sediment test result without Loeb criteria, and the suspected site of infection at ED. The selection of these variables is not unexpected because accurate diagnoses among older adults are complicated by their frequently atypical presentation. Nonspecific symptoms, such as altered mental status or malaise, are more prevalent in older patients presenting to the ED and may further increase diagnostic uncertainty.²⁶ Furthermore, patients with dementia may have difficulty expressing their symptoms. Finally, distinguishing real urinary tract infections from asymptomatic bacteriuria can be challenging, as demonstrated by Flokas et al,²⁷ who showed that the presence of nonspecific signs and symptoms and overreliance on laboratory tests may explain the excessive misinterpretation of asymptomatic bacteriuria as being a urinary tract infection. The C statistic of 0.69 indicates a moderate performance and hence will probably not directly benefit the clinician. However, we report this as a starting point for future research.

Strengths and Limitations

A strength of this study is its comprehensiveness, including all adult patients from whom a blood culture was taken at an ED of a large hospital during a 1-year period, reflecting common daily practice. This study also has some limitations. First, this study involves electronic records and may therefore be subject to information bias. Second, the study only looked at differences between sites of infection, but a wrong diagnosis within the same site of infection (eg, diverticulitis and cholecystitis) may also have implications. Third, we have not assessed at what time during a hospital stay the diagnosis has changed and how this might affect the results. Fourth, we do not know whether prescribed antibiotics were appropriate for the actual infection.

CONCLUSIONS

In this cohort study, misdiagnosed site of infection occurred in 1 of 9 patients hospitalized for suspected bacteremia but was not associated with worse short-term clinical outcomes. This finding might be explained by adherence to sepsis guidelines. Risk factors for misdiagnosed site of infection include older age, dementia, a positive urine sediment test result without urinary symptoms, and particular suspected sites of infection (intravascular, CNS, and bone and joint). Additional studies are needed to confirm these risk factors for misdiagnosis of infection at the ED and to further explore the association of misdiagnosed site of infection and adverse outcomes.

AUTHOR CONTRIBUTIONS

Ms Dregmans and Dr Kaal had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Ms Dregmans and Dr Kaal contributed equally to this study.

Concept and design Dregmans, Kaal, Meziyerh, Steyerberg, van Nieuwkoop. **Acquisition, analysis, or interpretation of data:** All authors. **Drafting of the manuscript:** Dregmans, Kaal, Meziyerh, van Nieuwkoop. **Critical revision of the manuscript for important intellectual content:** All authors. **Statistical analysis:** Dregmans, Kaal, Meziyerh, Steyerberg, van Nieuwkoop. **Obtained funding:** Meziyerh, van Nieuwkoop. **Administrative, technical, or material support:** Kaal, Meziyerh. **Supervision:** Kaal, Meziyerh, van Nieuwkoop.

CONFLICT OF INTEREST DISCLOSURES

None reported.

FUNDING/SUPPORT

This work was supported by unrestricted grant T18-040 from the Haga Scientific Research Fund.

ROLE OF THE FUNDER/SPONSOR

The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

SUPPLEMENTARY MATERIAL



REFERENCES

1. Guly HR. Diagnostic errors in an accident and emergency department. *Emerg Med J*. 2001;18(4):263-269. doi: 10.1136/emj.18.4.263
2. Hussain F, Cooper A, Carson-Stevens A, et al. Diagnostic error in the emergency department: learning from national patient safety incident report analysis. *BMC Emerg Med*. 2019;19(1):77. doi:10.1186/s12873-019-0289-3
3. Jayaprakash N, Chae J, Sabov M, Samavedam S, Gajic O, Pickering BW. Improving diagnostic fidelity: an approach to standardizing the process in patients with emerging critical illness. *Mayo Clin Proc Innov Qual Outcomes*. 2019;3(3):327-334. doi:10.1016/j.mayocpiqo.2019.06.001
4. Newman-Toker DE, Wang Z, Zhu Y, et al. Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the "Big Three". *Diagnosis (Berl)*. 2020;8(1):67-84. doi:10.1515/dx-2019-0104
5. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf*. 2013;22(suppl 2):ii21-ii27. doi:10.1136/bmjqs-2012-001615
6. Newman-Toker DE. A unified conceptual model for diagnostic errors: underdiagnosis, overdiagnosis, and misdiagnosis. *Diagnosis (Berl)*. 2014;1(1):43-48. doi:10.1515/dx-2013-0027
7. Zwaan L, Singh H. The challenges in defining and measuring diagnostic error. *Diagnosis (Berl)*. 2015;2(2): 97-103. doi:10.1515/dx-2014-0069
8. Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, The National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. In: Balogh EP, Miller BT, Ball JR, eds. *Improving Diagnosis in Health Care*. National Academies Press; 2015.
9. Zwaan L, de Bruijne M, Wagner C, et al. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med*. 2010;170(12):1015-1021. doi:10.1001/archinternmed.2010.146
10. Hautz WE, Kämmer JE, Hautz SC, et al. Diagnostic error increases mortality and length of hospital stay in patients presenting through the emergency room. *Scand J Trauma Resusc Emerg Med*. 2019;27(1):54. doi:10.1186/s13049-019-0629-z
11. Atamna A, Shiber S, Yassin M, Drescher MJ, Bishara J. The accuracy of a diagnosis of pneumonia in the emergency department. *Int J Infect Dis*. 2019;89:62-65. doi:10.1016/j.ijid.2019.08.027
12. Abe T, Tokuda Y, Shiraishi A, et al; JAAM SPICE Study Group. In-hospital mortality associated with the misdiagnosis or unidentified site of infection at admission. *Crit Care*. 2019;23(1):202. doi:10.1186/s13054-019-2475-9
13. Newman-Toker DE, Schaffer AC, Yu-Moe CW, et al. Serious misdiagnosis-related harms in malpractice claims: The "Big Three"—vascular events, infections, and cancers. *Diagnosis (Berl)*. 2019;6(3):227-240. doi:10.1515/dx-2019-0019
14. Caterino JM, Stevenson KB. Disagreement between emergency physician and inpatient physician diagnosis of infection in older adults admitted from the emergency department. *Acad Emerg Med*. 2012;19(8):908-915. doi:10.1111/j.1553-2712.2012.01415.x

15. Morr M, Lukasz A, Rübig E, Pavenstädt H, Kümpers P. Sepsis recognition in the emergency department: impact on quality of care and outcome? *BMC Emerg Med.* 2017;17(1):11. doi:10.1186/s12873-017-0122-9
16. Tellor B, Skrupky LP, Symons W, High E, Micek ST, Mazuski JE. Inadequate source control and inappropriate antibiotics are key determinants of mortality in patients with intra-abdominal sepsis and associated bacteremia. *Surg Infect (Larchmt).* 2015;16(6):785-793. doi:10.1089/sur.2014.166
17. Netherlands Trial Register. Predicting Microbiological and Clinical Outcomes of Patients With Suspected Bacteremia Presenting at Emergency Department: A Prospective Observational Study. NL7852. Accessed May 12, 2022. <https://www.trialregister.nl/trial/7852>
18. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA.* 2016;315(8):801-810. doi:10.1001/jama.2016.0287
19. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol.* 2011;173(6):676-682. doi:10.1093/aje/kwq433
20. Loeb M, Bentley DW, Bradley S, et al. Development of minimum criteria for the initiation of antibiotics in residents of long-term-care facilities: results of a consensus conference. *Infect Control Hosp Epidemiol.* 2001;22 (2):120-124. doi:10.1086/501875
21. Dutch Working Party on Antibiotic Policy (SWAB). SWAB Guidelines for Antibacterial Therapy of Adult Patients With Sepsis. SWAB; 2010. Accessed May 12, 2022. <https://swab.nl/exec/file/download/44>
22. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B.* 1996;58(1):267-288. doi:10.1111/j.2517-6161.1996.tb02080.x
23. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res.* 2011;20(1):40-49. doi:10.1002/mpr.329
24. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* Springer; 2009. doi:10.1007/978-0-387-77244-8
25. World Health Organization. *Global Action Plan on Antimicrobial Resistance.* World Health Organization; 2015.
26. Caterino JM, Kline DM, Leininger R, et al. Nonspecific symptoms lack diagnostic accuracy for infection in older patients in the emergency department. *J Am Geriatr Soc.* 2019;67(3):484-492. doi:10.1111/jgs.15679
27. Flokas ME, Andreatos N, Alevizakos M, Kalbasi A, Onur P, Mylonakis E. Inappropriate management of asymptomatic patients with positive urine cultures: a systematic review and meta-analysis. *Open Forum Infect Dis.* 2017;4(4):ofx207. doi:10.1093/ofid/ofx207



Chapter 6

Outcomes of community-acquired pneumonia using the Pneumonia Severity Index versus the CURB-65 in routine practice of emergency departments

Anna G. Kaal, Linde op de Hoek, Davinia T. Hochheimer, Corline Brouwers, Joost W. Wiersinga, Dominic Snijders, Katrijn L. Rensing, Christel E. van Dijk, Ewout W. Steyerberg, Cees van Nieuwkoop

ABSTRACT

Background

The Pneumonia Severity Index (PSI) and the CURB-65 score assess disease severity in patients with community acquired pneumonia (CAP). We compared the clinical performance of both prognostic scores according to clinical outcomes and admission rates.

Methods

A nationwide retrospective cohort study was conducted using claims data from adult CAP patients presenting to the emergency department (ED) in 2018 and 2019. Dutch hospitals were divided into three categories: 'CURB-65 hospitals' (n=25), 'PSI hospitals' (n=19) and hospitals using both ('no-consensus hospitals', n=15). Main outcomes were hospital admission rates, intensive care unit admissions, length of hospital stay, delayed admissions, readmissions and all-cause 30-day mortality. Multilevel logistic and Poisson regression analysis were used to adjust for potential confounders.

Findings

Of 50,984 included CAP patients, 21,157 were treated in CURB-65 hospitals, 17,279 in PSI hospitals and 12,548 in no-consensus hospitals. The 30-day mortality was significantly lower in CURB-65 hospitals versus PSI hospitals (8.6% and 9.7%, adjusted odds ratio (aOR) 0.89, 95% CI: 0.83-0.96, $p = 0.003$). Other clinical outcomes were similar between CURB-65 hospitals and PSI hospitals. No-consensus hospitals had higher admission rates compared to the CURB-65 and PSI hospitals combined (78.4% and 81.5%, aOR 0.78, 95% CI: 0.62-0.99).

Interpretation

In this study, using the CURB-65 in CAP patients at the ED is associated with similar and possibly even better clinical outcomes compared to using the PSI. After confirmation in prospective studies, the CURB-65 may be recommended over the use of the PSI since it is associated with lower 30-day mortality and more user-friendly.

INTRODUCTION

Community acquired pneumonia (CAP) is one of the most common infectious diseases, causing substantial clinical and economic burden worldwide.¹⁻³ To guide physicians in estimating disease severity in CAP patients, prediction tools such as the Pneumonia Severity Index (PSI) and the CURB-65 score have been developed.^{4,5} The PSI is a prognostic score based on 20 variables including demographics, comorbidities, and clinical variables. The CURB-65 consists of five elements: confusion, uraemia, respiratory rate, blood pressure, and age ≥ 65 years.^{4,5} The PSI was developed to identify patients with a low 30-day mortality risk, and is considered a safe way to determine need for hospitalization.^{4,6-9} While the CURB-65 was developed to stratify patients into mortality risk groups with separate management strategies, no studies investigated the effect of the CURB-65 on admission rates.⁵ The prognostic performance of both scores is considered to be similar by most studies, although some found the PSI to perform better in predicting mortality and identifying low risk patients.¹⁰⁻¹³

International guidelines have yet to reach consensus on which prediction tool to prefer for clinical practice and how to use it. The British Thoracic Society and NICE guidelines recommend using the CURB-65 for severity assessment, and for deciding on treatment strategy and treatment location in adults with CAP.^{14,15} Contrarily, the American Thoracic Society and Infectious Diseases Society of America (IDSA) recommends the PSI as a decision aid only to determine treatment location, while stating that there is a paucity of evidence regarding the effectiveness and safety of the CURB-65.¹⁶

The Dutch CAP guideline recommends using a prognostic tool to assess disease severity and to decide on treatment strategy.¹⁷ Importantly, the guideline does not prefer either the PSI or the CURB-65. This gave us the unique opportunity to compare the clinical performance of both prognostic scores on a large scale, by comparing clinical outcomes of CAP patients treated at hospitals that use the PSI to those treated at hospitals that use the CURB-65. The aim of this study is to evaluate the PSI and the CURB-65 in clinical practice, by comparing clinical outcomes and admission rates of CAP patients presenting to the emergency department (ED).

METHODS

Setting and data sources

We conducted a retrospective study using claims data provided by the centre for information of Dutch health insurers, Vektis, to the National Health Care Institute (NHCI) of the Netherlands. Vektis collects data from all health insurers which include,

among other, insurance claims of medical specialist care and medication claims of pharmacies.

Claims data

In the Dutch healthcare system, hospital care is financed through Diagnosis Treatment Combinations (DTCs). All patients receiving hospital care are assigned a DTC according to their diagnosis. This DTC covers all hospital activities and interventions that a patient with that diagnosis might receive, and is specific for both diagnosis and medical specialty (e.g. DTC pneumonia, department of internal medicine).

Patients

All patients aged ≥ 18 years with a pneumonia DTC started between January 1st 2018 and December 31st 2019, with primary involvement of internal medicine or pulmonology, were selected. Of those, only patients with an ED visit, including the performance of chest imaging (X-ray or Computed Tomography) at the start of the DTC were included, as that would be the setting in which the PSI and CURB-65 are used. If a patient had more than one episode of pneumonia, only the first was included. To restrict to community acquired cases, patients who had been discharged from a hospital up to 30 days prior to the DTC registration were excluded.

Variables and outcomes

Demographic information including age, sex, and comorbidities was available in the claims database. A neighbourhood socioeconomic status (SES) from 2016 was available through Netherlands Bureau for Economic Policy Analysis (CPB). This is measured using status scores indicating the relative status of a neighbourhood. It is derived from education, income and position on the labour market and was divided into quartiles for analysis (a higher quartile indicating a higher SES).¹⁸ Comorbidities were identified through additional DTC registrations. DTCs started within five years prior to the pneumonia DTC were included to include current comorbidities. These DTCs were classified into larger categories, based on the categories in the PSI as defined by Fine et al: neoplastic disease, liver disease, congestive heart failure, cerebrovascular disease and chronic renal disease.⁴ In addition, we assessed the presence of cardiovascular disease, pulmonary disease, diabetes mellitus, immunosuppression, neurologic disease and HIV-infection. Smoking status or body mass index was not available.

By assessing DTCs only, information on relevant diseases treated by a general practitioner could be missed. Therefore, chronic use of immunosuppressants and

diabetes medication in the year prior to the pneumonia was identified through medication claims of pharmacies. Patients were considered immunocompromised if they used at least one immunosuppressant in combination with a DTC suggesting impaired immunity. The diabetes group was compiled by patients either using diabetes medication, having a diabetes DTC, or both.

The main outcomes of this study were hospital admission rate, readmissions, ICU admissions, delayed admissions, all-cause 30-day mortality and length of hospital stay. Readmission was defined as a discharge from the hospital followed by admission within 7 days. Delayed admission was defined as initial outpatient treatment, followed by admission to the hospital within 7 days. More detailed definitions, an overview of DTCs per category and included medication can be found in the appendix.

Hospitals

The NHCI conducted a survey in March 2020 among all pulmonologists and internal medicine specialists in the Netherlands asking which prognostic score for CAP was used in their hospital during 2018-2019 and how this prognostic score was used.¹⁹ Response from at least one physician from both departments was required. If this was not the case, a physician on call from the missing specialty was contacted by phone. We received a response from 93% of all hospitals.

After collecting all survey results and answers received by phone, hospitals were divided into 'PSI hospitals', 'CURB-65 hospitals' and 'no-consensus hospitals'. This last category consists of hospitals using both prognostic scores within the same department or different prognostic scores between internal medicine and pulmonology departments. As claims data were hospital specific, we were able to specify which prognostic score was used for severity assessment in the selected patients.

Statistical analysis

For each outcome, patients from 'PSI hospitals' were compared to patients from 'CURB-65 hospitals' and 'no-consensus hospitals'. In addition, patients from 'no-consensus hospitals' were compared to patients from PSI and CURB-65 hospitals as one group ('consensus'), to examine differences between hospitals that use one prognostic score and hospitals that use both.

Crude and adjusted odds ratios (ORs) with 95%-confidence intervals (95%-CI) were calculated for all outcomes. Multilevel logistic regression analyses (iterative generalized least squares (IGLS), 2nd order PQL, random intercept at hospital level)

were performed for dichotomous outcomes. A multilevel Poisson linear regression analysis (IGLS, 2nd order PQL, random intercept at hospital level) was performed for length of stay, with a correction for overdispersion. We corrected for possible confounders: age, sex, neighbourhood SES, comorbidities, specialism (internal medicine or pulmonology), and type of hospital (general or teaching hospital). Furthermore, a variable was added reflecting the number of days that patients were alive in the investigated period to account for time-dependent risk, except for the analysis of 30-day mortality. For 30-day mortality, three multilevel regression subgroup analyses were performed based on predefined age categories, admission (yes/no) and ICU admission (yes/no). P-values of < 0.05 were considered statistically significant. SAS Enterprise Guide 7.1 was used for the descriptive statistics. MLwiN 3.05 was used for the multilevel analyses.

RESULTS

Hospitals

A total of 59,059 patients from 73 Dutch hospitals met the eligibility criteria (Figure 1). Patients from the eight university hospitals were excluded as they were all either CURB-65 hospitals or no-consensus hospitals, making a comparison to university PSI hospitals impossible. One hospital used neither PSI or CURB-65 and was therefore excluded. Five other hospitals were excluded, as we were unable to determine which prognostic score was used. Finally, 257 patients from whom the zip code and therefore the SES was not available were excluded. A total of 50,984 patients from 59 hospitals were included in the final analysis (Figure 1).

Questionnaire

The respondents to the questionnaire estimated that in 79% of patients a prognostic score was used. 83% of the respondents used the prognostic score to decide on antibiotic treatment, 49% used it to determine the need for hospitalization, and 28% used it to decide whether a patient should be treated at an ICU.¹⁹

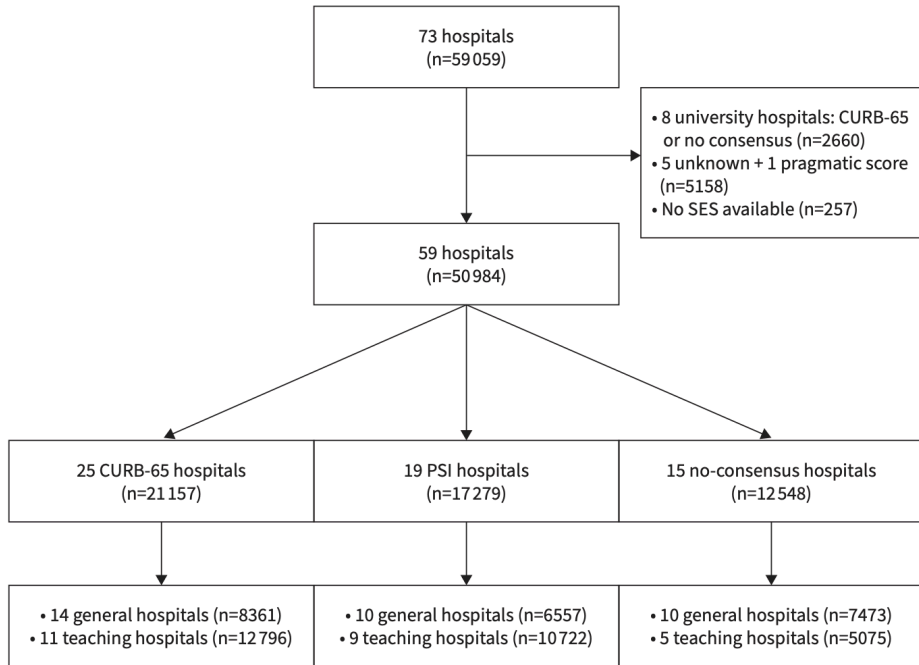


Figure 1 Flowchart of hospital and patient selection. PSI: Pneumonia Severity Index; SES: socio-economic status.

Table 1. Baseline patient characteristics

	CURB-65 hospital	PSI hospital	No-consensus hospital	Total
Total patients	21 157 (41.5)	17,279 (33.9)	12,548 (24.6)	50,984 (100)
Female patients	10,096 (47.7)	8,019 (46.4)	5,947 (47.4)	24,062 (47.2)
Age years	72 (60-81)	72 (60-81)	72 (61-82)	72 (60-81)
Age range years				
18-49 years, N (%)	2,838 (13.4)	2,184 (12.6)	1,493 (11.9)	6,515 (12.8)
50-64 years, N (%)	4,185 (19.8)	3,469 (20.1)	2,458 (19.6)	10,112 (19.8)
65-74 years, N (%)	5,083 (24.0)	4,203 (24.3)	3,008 (24.0)	12,294 (24.1)
75-84 years, N (%)	5,500 (26.0)	4,471 (25.9)	3,339 (26.6)	13,310 (26.1)
> 85 years, N (%)	3,551 (16.8)	2,952 (17.1)	2,250 (17.9)	8,753 (17.2)
SES score				
Quartile 1, N (%)	5,020 (23.7)	6,219 (36.0)	3,299 (26.3)	14,538 (28.5)
Quartile 2, N (%)	6,011 (28.4)	4,166 (24.1)	3,716 (29.6)	13,893 (27.2)
Quartile 3, N (%)	5,533 (26.2)	3,861 (22.4)	2,973 (23.7)	12,367 (24.3)
Quartile 4, N (%)	4,593 (21.7)	3,033 (17.6)	2,560 (20.4)	10,186 (20.0)
Comorbidities				
History of neoplastic disease ^a	3,849 (18.2)	3,212 (18.6)	2,192 (17.5)	9,253 (18.1)
Liver disease ^a	292 (1.4)	229 (1.3)	134 (1.1)	655 (1.3)
Congestive heart failure ^a	1,993 (9.4)	1,764 (10.2)	1,279 (10.2)	5,036 (9.9)
Cerebrovascular disease ^a	1,574 (7.4)	1,285 (7.4)	923 (7.4)	3,782 (7.4)
Chronic renal disease ^a	1,153 (5.5)	997 (5.8)	635 (5.1)	2,785 (5.5)
Cardiovascular disease [¶]	4,250 (20.1)	3,716 (21.5)	2,559 (20.4)	10,525 (20.6)
Pulmonary disease ^c	4,048 (19.1)	3,455 (20.0)	2,468 (19.7)	9,971 (19.6)
Diabetes mellitus ^d	3,713 (17.6)	3,278 (19.0)	2,285 (18.2)	9,276 (18.2)
Immunocompromised ^e	576 (2.7)	515 (3.0)	367 (2.9)	1,458 (2.9)
Neurologic disease ^f	457 (2.2)	403 (2.3)	288 (2.3)	1,148 (2.3)
HIV infection [¶]	66 (0.3)	32 (0.2)	15 (0.1)	113 (0.2)
Treated at pulmonology department	15,051 (71.1)	12,387 (71.7)	8,905 (71.0)	36,343 (71.3)
Treated at teaching hospital	12,796 (60.5)	10,722 (62.1)	5,075 (40.4)	28,593 (56.1)
Main clinical outcomes				
Admission	16,331 (77.2)	13,807 (79.9)	10,230 (81.5)	40,368 (79.2)
ICU admission	1,869 (8.8)	1,395 (8.1)	1,077 (8.6)	4,341 (8.5)
30-day mortality	1,827 (8.6)	1,682 (9.7)	1,119 (8.9)	4,628 (9.1)

^a As defined by Fine et al. (1997)⁴[¶] See appendix for the specific DTC codes^c DTC code COPD, asthma or cystic fibrosis^d DTC code diabetes mellitus type 1 or type 2 or the use of at least one type of diabetes medication in the year prior to the pneumonia DTC, or both^e DTC code of transplantation recipients or auto-immune disease (see appendix), and the use of at least one immunosuppressant in the year prior to the pneumonia DTC^f DTC code Parkinson's disease, ALS, myasthenia gravis, multiple sclerosis or other neuromuscular disease

DTC: Diagnosis Treatment Combination healthcare product; PSI: Pneumonia Severity Index

Baseline characteristics

24.062 patients were female (47%), the median age was 72 years (Table 1). The most prevalent comorbidities across all 50.984 patients were cardiovascular disease (10.525, 21%), pulmonary disease (9.971, 20%), neoplastic disease (9.253, 18%) and diabetes mellitus (9.276, 18%). The patients from PSI hospitals were more likely to have a SES in the lowest quartile compared to the other patients. Overall, 36.343 out of 50.984 patients were managed by the pulmonology department (71%). The majority of the consensus hospitals classified as teaching hospitals; 61% of CURB-65 hospitals and 62% of PSI hospitals. In contrast, most of the no-consensus hospitals were general hospitals (60%). (Table 1)

OUTCOMES

Overall, 40.368 out of 50.984 (79.2%) patients were admitted to the hospital directly after their ED visit and 4.341 out of 50.984 (8.5%) were admitted to the ICU. The all-cause 30-day mortality was 9.1% (4.628/50.984, Table 1). Out of 10.616 patients primarily treated as outpatients, 832 (7.8%) were admitted within seven days after presentation at the ED. Of all 41.200 combined admissions and delayed admissions, 977 patients (2.4%) were readmitted within seven days after discharge.

CURB-65 versus PSI

The 30-day mortality was significantly lower in the CURB-65 hospitals compared to the PSI hospitals; 8.6% vs 9.7% (Table 2). This association was statistically significant with an adjusted odds ratio (aOR) of 0.89, 95% CI: 0.83-0.96, $p=0.003$. There was a trend towards lower admission rates in the CURB-65 hospitals than in the PSI hospitals; 77.2% vs 79.9% (aOR 0.81, 95% CI: 0.64-1.02). ICU admission rates, readmissions and delayed admissions were similar between the CURB-65 and PSI hospitals (Table 2). Lastly, for all admitted patients, the median length of stay was 5 days (IQR 4-8) for the CURB-65 hospitals and 6 days (IQR 4-8) for the PSI hospitals, with an adjusted RR (aRR) of 0.99 (95% CI: 0.94-1.04).

Table 2. Multilevel analysis of outcomes of patients with community-acquired pneumonia using CURB-65 versus PSI

	CURB-65 (N = 21157)	PSI (N = 17279)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Admission, N (%)	16,331 (77.2)	13,807 (79.9)	0.81 (0.64-1.02)	0.81 (0.64-1.02)
Readmission ^a , N (%)	410 (2.5)	329 (2.3)	1.03 (0.86-1.24)	1.02 (0.85-1.22)
Delayed admission ^b , % N (%)	371 (7.7)	275 (7.9)	0.98 (0.79-1.21)	0.99 (0.80-1.22)
ICU admission ^c , N (%)	1,869 (8.8)	1,395 (8.1)	1.10 (0.93-1.31)	1.14 (0.96-1.35)
30-day mortality ^c , % N (%)	1,827 (8.6)	1,682 (9.7)	0.88 (0.81-0.96)	0.89 (0.83-0.96)

a. Percentage out of the total number of (delayed) admissions

b. Percentage out of the total number of initial outpatients

c. Percentage out of the total number of patients

Readmission was defined as patients who were discharged from the hospital but readmitted within 7 days. Delayed admission was defined as ED pneumonia patients initially treated as outpatients, but who were admitted to the hospital within 7 days. PSI: Pneumonia Severity Index

Consensus versus no-consensus

Admission rates were significantly lower in consensus hospitals compared to no-consensus hospitals: 78.4% vs 81.5% (aOR 0.78, 95% CI: 0.62-0.99, Table 3). Other outcomes were all similar, including the median length of stay (aRR 1.03, 95% CI: 0.98-1.08).

Table 3. Multilevel analysis of outcomes of patients with community-acquired pneumonia using either CURB-65 or PSI versus no-consensus hospitals

	Consensus (PSI or CURB-65) (N = 38,436)	No consensus (N = 12548)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Admission, N (%)	30,138 (78.4)	10,230 (81.5)	0.79 (0.62-0.999)	0.78 (0.62-0.99)
Readmission ^a , N (%)	739 (2.4)	238 (2.3)	1.06 (0.88-1.26)	1.05 (0.87-1.25)
Delayed admission ^b , N (%)	646 (7.8)	186 (8.0)	0.98 (0.79-1.21)	0.99 (0.80-1.23)
ICU admission ^c , N (%)	3,264 (8.5)	1,077 (8.6)	0.96 (0.81-1.13)	0.96 (0.81-1.13)
30-day mortality ^c , N (%)	3,509 (9.1)	1,119 (8.9)	1.05 (0.96-1.16)	1.07 (0.98-1.16)

^a Percentage out of the total number of (delayed) admissions

^b Percentage out of the total number of outpatients

^c Percentage out of the total number of patients

Consensus was defined as a hospital where both the internal and pulmonology department reported the use of the same prognostic score (Either PSI or CURB-65). Readmission was defined as patients who were discharged from the hospital but readmitted within 7 days. Delayed admission was defined as patients initially treated as outpatients, but who were admitted to the hospital within 7 days.

Subgroup analyses

In the analysis by age, most subgroups showed a decrease in mortality when treated in a CURB-65 hospital versus a PSI hospital (Fig 2). There was a stronger association in the age group of 50–64 (aOR 0.71 (0.55–0.92)). In the age group 18 to 49, no reliable estimate could be made due to the low mortality rate. The subgroup analysis by admission showed similar associations for admitted patients (aOR 0.90 (0.83–0.97)) and for those treated as outpatients (aOR 0.91 (0.72–1.15)). Finally, for both patients with and without ICU admission, the CURB-65 was associated with a decrease in 30-day mortality (Fig 2).

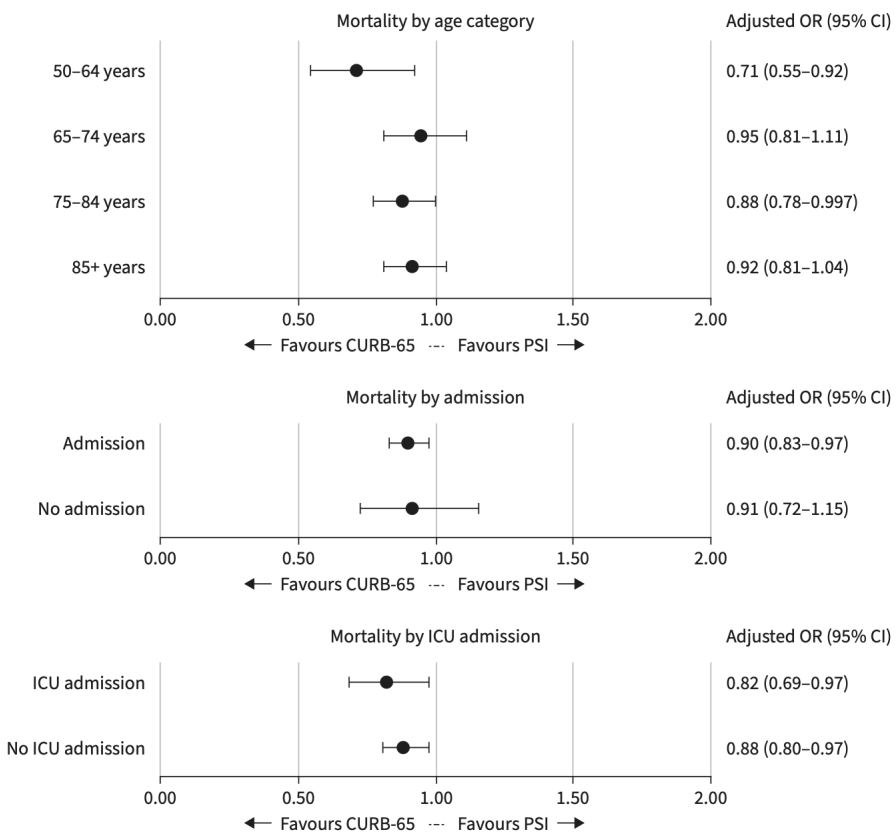


Figure 2. Subgroup analysis for the primary outcome 30-day mortality in patients with community-acquired pneumonia: CURB-65 versus Pneumonia Severity Index (PSI). ICU: intensive care unit.

DISCUSSION

We found that using the CURB-65 as a severity assessment tool among ED CAP patients was associated with a lower risk of 30-day mortality compared to the PSI. In addition, the admission rates were higher among hospitals that did not consistently use only one of the prognostic scores in the management of CAP. Based on this large multicentre study, using the CURB-65 in CAP patients at the ED leads to similar and possibly even better clinical outcomes compared to the PSI.

The overall 30-day mortality rate in this study was 9.1%, slightly higher than mortality rates in previous Dutch studies of 6.7% and 6.8%.^{20,21} This difference could be explained by the fact that the patients in these studies were on average younger with less comorbidities. Our mortality rate was in line with the PSI study by Fine et al., and the CURB-65 study by Lim et al., showing mortality rates of 10.2% and 9%.^{4,5}

It is a striking finding that the CURB-65 is associated with a lower 30-day mortality compared to the PSI. There are several potential explanations for this observation. Several studies showed that the CURB-65 classifies more patients as 'severe CAP' (score 3-5) compared to the PSI (score of 5).^{12,20,22} This finding, combined with the fact that the Dutch CAP guideline recommends using the CURB-65 or the PSI for deciding upon antibiotic treatment, could potentially lead to different empirical treatment in CURB-65 versus PSI hospitals.¹⁷ We do not have data on antibiotic use, however, since 83% of respondents use the prognostic score to support their choice of antibiotics, it seems realistic to think that there has been a difference in treatment between the CURB-65 and PSI hospitals.¹⁹ In severe CAP, it is advised to start with a 2nd or 3rd generation cephalosporin intravenously, whereas in mild and moderate-severe CAP, narrow-spectrum beta-lactam antibiotics (including oral regimens) are recommended. Thus, categorizing more patients as 'severe CAP' when using CURB-65, would likely have resulted in more patients being treated with cephalosporins, which might explain the lower mortality rate. The fact that the difference was especially visible in the age category of 50 to 65 years would also support the hypothesis, since the PSI tends to underestimate disease severity for younger patients.²³ Another explanation could be that, since more patients are judged as severe CAP when using CURB-65, more patients received intensive monitoring and associated supportive care measures resulting in better outcome. One could hypothesize that patients in CURB-65 hospitals were less ill, especially since they were less often admitted. However, the analysis of only admitted patients showed similar results. In another analysis, we included the academic hospitals which were seven CURB-65 hospitals and one no-consensus hospital. One would expect that academic hospitals treat patients with a higher risk of adverse outcomes. However, this also did not change our results (See appendix). Of course, it is possible

that residual confounding accounts for the observed difference. Notably, the PSI hospitals treated patients with relatively lower SES. Although analyses were corrected for SES as a confounder, low SES is for instance associated with smoking, something that we could not take into account.²⁴ Finally, it should be noted that international guidelines differ in whether prognostic scores should be used to guide antibiotic treatment or (ICU) admission policy. Such differences should be kept in mind when interpreting mortality outcomes in an international setting.

Another interesting finding is that we found lower admission rates in consensus hospitals compared to no-consensus hospitals, without an increase in adverse outcomes. This suggests that not consistently using a prognostic score within one hospital could lead to suboptimal admission policy. It has been shown that using the PSI is associated with lower admission rates without compromising safety outcomes.⁶⁻⁹ To our knowledge, there are currently no data regarding the CURB-65 and its association with hospitalization rates. In this study, the admission rates in the CURB-65 hospitals were similar to the PSI hospitals, with a trend for even lower admission rates (77.2% versus 79.9%, OR 0.81 (0.64-1.02)). This could be explained by the fact that the CURB-65 does not only classify more patient in the high risk categories, but also classifies more patients into the lower risk categories when compared to the PSI.^{12,17,20}

The major strength of this study is the large study population, representing the whole population across a Western country. Another strength is the inclusion of patients from two years, limiting seasonal effects. Moreover, only patients from 2018 and 2019 were included, i.e. before the COVID-19 epidemic. However, using claims data inevitably leads to important limitations. First, we did not have access to information at the individual level such as received treatment, diagnostic results, resuscitation codes and the actual patient's risk score. However, it is not to be expected that the patient populations would be significantly different across the three hospital categories. Second, we have limited information on guideline adherence other than the questionnaire results. Our hypotheses to explain the observed differences are based on the assumption that the patient's risk classification leads to a specific treatment. Although it is known that guideline adherence is suboptimal, we do not have a reason to assume that guideline adherence would be different between the PSI and the CURB-65 hospitals.²⁰ Furthermore, the questionnaire results suggest that the majority of respondents do use the prognostic score to decide on treatment. Third, we have limited information about differences between hospitals, such as differences in local treatment guidelines, different hospital facilities or differences in expertise of treating physicians. However, the difference in quality of care between Dutch hospitals is generally considered to be minimal. Furthermore, we corrected

all analysis for type of hospital and all hospital categories were evenly distributed across the different geographic regions. Finally, we tried to categorize each hospital as accurately as possible. Although unlikely, it cannot be fully ruled out that a hospital was misclassified.

To our knowledge, this is the largest study comparing clinical outcomes when using the PSI or CURB-65 in clinical practice of adult CAP patients. Previous studies have mainly focused on the predictive abilities, with most studies finding similar results for the PSI and the CURB-65, while our study focused on clinical outcomes.^{10,11} Although claims data inevitably has its limitations, we believe that this study shows that consistently using the PSI or CURB-65 in one hospital is associated with lower admission rates without increasing adverse outcomes. Furthermore, using the CURB-65 in CAP patients at the ED is associated with similar and possibly even better clinical outcomes compared to the PSI. Since the CURB-65 is also more user-friendly than the PSI as it only contains 5 variables, we conclude that after confirmation of our findings in well conducted prospective studies, the CURB-65 may be recommended over the use of the PSI.

Ethics

According to Dutch legislation, there was no need for approval by a medical ethics committee since we used pseudo-anonymized claims data.

Data statement

Due to privacy legislation, underlying claims data cannot be made available.

Declaration of interests

JW reports two grants, the Vidi grant from the Dutch Research Council and a grant on COVID from The Netherlands Organisation for Health Research and Development. Furthermore, his host institution received ad hoc consultancy fees for GSK (DSMB), Pfizer, AstraZeneca and Sobi. All other authors report no conflicts of interest.

Contributors' statement

Conceptualization: AK, LH, DH, CB, WJW, DS, KR, CN. Data curation: AK, LH, DH, CD. Software: AK, DH, CD. Formal analysis: AK, LH, DH, CD. Supervision: ES, CN. Validation: All authors. Investigation: AK, LH, DH, CD. Visualisation: AK, LH. Methodology: AK, LH, DH, CB, KR, CD, CN. Writing - Original draft: AK, LH. Writing - Review and editing: All authors. All authors had access to the data, the underlying data was verified by AK, LH, DH and CD.

SUPPLEMENTAL MATERIAL



REFERENCES

1. Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax*. 2012;**67**:71–79. doi:10.1136/thx.2009.129502
2. McLaughlin JM, Khan FL, Thoburn EA, Isturiz RE, Swerdlow DL. Rates of hospitalization for community-acquired pneumonia among US adults: A systematic review. *Vaccine*. 2020;**38**(4):741–751. doi:10.1016/j.vaccine.2019.10.101
3. Torres A, Cilloniz C, Niederman MS, et al. Pneumonia. *Nature Reviews Disease Primers*. 2021;**7**:25. doi: 10.1038/s41572-021-00259-0
4. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997;**336**:243–50. doi:10.1056/NEJM199701233360402
5. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003;**58**:377–382. doi:10.1136/thorax.58.5.377
6. Chalmers JD, Akram AR, Hill AT. Increasing outpatient treatment of mild community-acquired pneumonia: Systematic review and meta-analysis. *European Respiratory Journal*. 2011;**37**:858–864. doi:10.1183/09031936.00065610
7. Renaud B, Coma E, Labarere J, et al. Routine use of the Pneumonia Severity Index for guiding the site-of-treatment decision of patients with pneumonia in the emergency department: a multicenter, prospective, observational, controlled cohort study. *Clin Infect Dis*. 2007;**44**:41–49. doi:10.1086/509331
8. Marrie TJ, Lau CY, Wheeler SL, Wong CJ, Vandervoort MK, Feagan BG. A controlled trial of a critical pathway for treatment of community-acquired pneumonia. *JAMA*. 2000;**283**:749–755.
9. Jo S, Kim K, Jung K, et al. The effects of incorporating a pneumonia severity index into the admission protocol for community-acquired pneumonia. *Journal of Emergency Medicine*. 2012;**42**:133–138. doi:10.1016/j.jemermed.2010.04.018
10. Chalmers JD, Singanayagam A, Akram AR, et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. *Thorax*. 2010;**65**:878–883. doi:10.1136/thx.2009.133280
11. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. *Thorax*. 2010;**65**:884–890. doi:10.1136/thx.2009.134072
12. Aujesky D, Auble TE, Yealy DM, et al. Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia. *American Journal of Medicine*. 2005;**118**:384–392. doi:10.1016/j.amjmed.2005.01.006
13. Ranzani OT, Prina E, Menéndez R, et al. New Sepsis Definition (Sepsis-3) and Community-acquired Pneumonia Mortality. A Validation and Clinical Decision-Making Study. *Am J Respir Crit Care Med*. 2017; **196**: 1287–1297, Nov 15, 2017
14. Lim WS, Baudouin SV, George RC, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*. 2009;**64**(Suppl III):iii1–iii55. doi:10.1136/thx.2009.121434

15. Pneumonia in adults: Diagnosis and management. National Institute for Health and Care Excellence (NICE); 2014. Last updated 16 September 2019.
16. Metlay JP, Waterer GW, Long AC, et al. Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med*. 2019;**200**:e45–e67. doi:10.1164/rccm.201908-1581ST
17. Wiersinga WJ, Bonten MJ, Boersma WG, et al. Management of community-acquired pneumonia in adults: 2016 guideline update from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). *Neth J Med*. 2018;**76**:4–13.
18. Knol F, Boelhouwer J, Veldheer V. Summary: Neighbourhood Status Development in the Netherlands 1998–2010. Den Haag: Sociaal en Cultureel Planbureau, 2012.
19. ZorgInstituut Nederland. Resultaten enquête risicoscores pneumonie. Maart 2020. <https://www.zorginstituutnederland.nl/publicaties/rapport/2020/03/01/resultaten-enquete-risicoscores-pneumonie>. Retrieval date 3rd of August 2022.
20. Huijts SM, Werkhoven CHV, Boersma WG, et al. Guideline adherence for empirical treatment of pneumonia and patient outcome. *The Netherlands Journal of Medicine*. 2013;**71**:502–507.
21. Raeven VM, Spoorenberg SMC, Boersma WG, et al. Atypical aetiology in patients hospitalised with community-acquired pneumonia is associated with age, gender and season; a data-analysis on four Dutch cohorts. *BMC Infect Dis*. 2016;**16**:299. doi:10.1186/s12879-016-1641-9
22. Chalmers JD, Rutherford J. Can we use severity assessment tools to increase outpatient management of community-acquired pneumonia? *Eur J Intern Med*. 2012;**23**:398–406. doi:10.1016/j.ejim.2011.10.002
23. Niederman MS. Making sense of scoring system in community acquired pneumonia. *Respirology*. 2009; **14**:327–335. doi: 10.1111/j.1440-1843.2009.01494.x
24. Barbeau EM, Krieger N, Soobader M. Working Class Matters: Socioeconomic Disadvantage, Race/Ethnicity, Gender, and Smoking in NHIS 2000. *Am J Public Health*. 2004;**94**:269–278. doi: 10.2105/ajph.94.2.269



Chapter 7

Diagnostic yield of bacteriological tests and predictors of severe outcome in adult patients with COVID-19 presenting to the emergency department

Anna Kaal, Lars Snel, Martijn Dane, Nathalie van Burgel, Thomas Ottens, Winifred Broekman, Lahssan El Bouazzaoui, Nikki Kolfschoten, Emile Schippers, Ewout Steyerberg, Soufian Meziyerh, Cees van Nieuwkoop

KEY MESSAGES

What is already known on this subject

- Reported rates of bacterial coinfection in patients with COVID-19 presenting at the emergency department (ED) are low (0%–8%). Nevertheless, guidelines advise to maximise efforts to collect blood cultures and other microbiological tests, in order to allow restrictive use of antibiotics.
- Early identification of patients with COVID-19 at risk for developing a severe disease course (intensive care unit admission and/or mortality) is important for clinical decision making. Several clinical and biochemical parameters are associated with a severe disease course. However, which parameters are of most value to identify patients at risk for severe disease course when presenting to the ED is unknown.

What this study adds

- In this retrospective study of patients receiving blood cultures at a hospital in The Netherlands, we found that 6.3% of patients with SARS-CoV-2 had positive blood cultures and all were considered contaminated. Overall bacterial coinfection using other diagnostics was 1.4%. This suggests that in patients known to have SARS-CoV-2, blood cultures are unlikely to be of use.
- Procalcitonin and the Pneumonia Severity Index seem most promising to recognise patients at risk for severe outcome of COVID-19 when presenting to the ED.

ABSTRACT

Background

Guidelines recommend maximal efforts to obtain blood and sputum cultures in patients with COVID-19, as bacterial coinfection is associated with worse outcomes. The aim of this study was to evaluate the yield of bacteriological tests, including blood and sputum cultures, and the association of multiple biomarkers and the Pneumonia Severity Index (PSI) with clinical and microbiological outcomes in patients with COVID-19 presenting to the emergency department (ED).

Methods

This is a substudy of a large observational cohort study (PredictED study). The PredictED included adult patients from whom a blood culture was drawn at the ED of Haga Teaching Hospital, The Netherlands. For this substudy, all patients who tested positive for SARS-CoV-2 by PCR in March and April 2020 were included. The primary outcome was the incidence of bacterial coinfection. We used logistic regression analysis for associations of procalcitonin, C reactive protein (CRP), ferritin, lymphocyte count and PSI score with a severe disease course, defined as intensive care unit admission and/or 30-day mortality. The area under the receiver operating characteristics curve (AUC) quantified the discriminatory performance.

Results

We included 142 SARS-CoV-2 positive patients. On presentation, the median duration of symptoms was 8 days. 41 (29%) patients had a severe disease course and 24 (17%) died within 30 days. The incidence of bacterial coinfection was 2/142 (1.4%). None of the blood cultures showed pathogen growth while 6.3% was contaminated. The AUCs for predicting severe disease were 0.76 (95% CI 0.68 to 0.84), 0.70 (0.61 to 0.79), 0.62 (0.51 to 0.74), 0.62 (0.51 to 0.72) and 0.72 (0.63 to 0.81) for procalcitonin, CRP, ferritin, lymphocyte count and PSI score, respectively.

Conclusion

Blood cultures appear to have limited value while procalcitonin and the PSI appear to be promising tools in helping physicians identify patients at risk for severe disease course in COVID-19 at presentation to the ED.

INTRODUCTION

Since the emergence of COVID-19, knowledge about the disease has quickly advanced. In times where large numbers of patients ask the most from healthcare facilities, it is of great relevance to identify those at high risk of developing a severe disease course. Currently, known risk factors for severe disease, defined as intensive care unit (ICU) admission and/or mortality, include older age and comorbidities such as diabetes, cardiovascular and pulmonary disease.^{1,2}

Furthermore, several biomarkers are associated with severe disease course in patients with COVID-19, including C-reactive protein (CRP), ferritin, lymphocyte count and procalcitonin (PCT).^{3,4} It is well-known that bacterial pulmonary coinfections are a major concern in respiratory viral infections. Zhou et al showed that in a cohort of non-survivors of COVID-19, 50% had a secondary infection.² Current data on patients with COVID-19 show that the incidence of bacterial coinfection on admission is relatively low, ranging from 0% to 8%.^{5,6} Differentiating between viral or bacterial pulmonary infection can be challenging. In this respect, PCT is a more accurate predictor than CRP in patients with community-acquired pneumonia (CAP) and influenza.^{7,8} Therefore, the American Thoracic Society & Infectious Diseases Society of America endorses withholding of antibiotics in patients with non-severe COVID-19 with low PCT levels.⁹ The Dutch Working Party on Antibiotic Policy (SWAB) advises a risk-based antibiotic policy based on the Pneumonia Severity Index (PSI) for patients with CAP.^{10,11} For patients with COVID-19, restrictive antibiotic use is encouraged except for those in the highest risk group and maximal efforts are recommended to obtain blood and sputum cultures.¹² It recognises that using PCT to guide antibiotic treatment might be a valid strategy, but evidence is lacking.^{11,12}

We were able to investigate a subset of patients with COVID-19 from an observational study, evaluating the value of PCT in patients visiting the emergency department (ED) from whom a blood culture was drawn. This gave us the opportunity to explore the previous reported association of PCT with severe course of COVID-19, and to investigate whether this might be explained by the presence of bacterial coinfection. We expected a relatively high rate of bacterial coinfection in this cohort since the performance of a blood culture is an indicator of suspected bacteraemia or sepsis. Besides exploring the role of bacterial coinfection, it is of interest to know which biomarkers are most useful to help physicians recognise patients with COVID-19 at high risk for severe disease when presenting to the ED.

We hypothesise that PCT is superior to more commonly used biomarkers such as CRP, ferritin and lymphocyte count in identifying patients with COVID-19 presenting to the ED with a bacterial coinfection and those who are at risk for a severe disease

course. The aim of this study is to explore the incidence of bacterial coinfection and the association of the PSI and multiple biomarkers with clinical and microbiological outcomes in patients with COVID-19 presenting to the ED.

METHODS

Study design and participants

This is a substudy of the PredictED study. The PredictED is a single-centre prospective observational study conducted at the ED of the Haga Teaching Hospital, a large hospital in The Hague, The Netherlands. Annually, around 50 000 patients present to this ED of which the majority are referred by a general practitioner. The PredictED study evaluates PCT as a marker for bacteraemia in patients presenting to the ED. All consecutive patients at the ED aged 18 years or older that had at least one blood culture taken at the treating physician's discretion were included in the PredictED. There were no exclusion criteria. Data collection of the PredictED was set from 1 April 2019 until 30 April 2020, which included the regional peak period of the first wave of the COVID-19 pandemic in the Netherlands during March and April 2020. Within this cohort, all patients with COVID-19 confirmed by a positive PCR test for SARS-CoV-2 were included in the present study. The PredictED study has been registered at [trialregister.nl](https://www.trialregister.nl) (NL7852; 7 July 2019).

Test methods

A SARS-CoV-2 infection was confirmed using an in-house real-time PCR using a combination of the QIA Symphony and ABI7500 on a nasopharyngeal/throat swab.¹³ Blood cultures were incubated using the BacTEC/Alert automated blood culture system. Determination was done with gram stain and with the Bruker Maldi-tof-MS biotyper. Due to the large amount of test requests in this period, standard incubation time was reduced from five to three days. Since PCT is not routinely measured at our ED, PCT was measured using leftover blood that was drawn for chemistry testing during the ED visit. PCT was measured with an automated rapid sensitive assay (Elecsys Brahms PCT on Roche Cobas E601, Basel, Switzerland, measuring range 0.02–100 ng/mL). The PCT result was not available to the attending physician. Other microbiological tests were performed at discretion of the treating physician. The reference tests were defined as the biomarkers that were measured in a routine test set for suspected patients with COVID-19 attending the ED, namely CRP, ferritin and lymphocyte count.

Definitions

Bacterial coinfection was defined as a positive PCR for *Mycoplasma pneumoniae* or *Chlamydia* spp, a positive pneumococcal or legionella urinary antigen test or bacterial growth in a blood or sputum culture recognised as a pathogen and not deemed a contaminant as reviewed by three members of our research team (AK, LS and SM). For blood cultures, Coagulase-negative staphylococci, *Bacillus* spp, *Corynebacterium* spp, *Cutibacterium* spp, Viridans group streptococci and *Lactobacillus* spp were considered contaminants unless associated with an intravascular catheter or device, suspected endocarditis or the presence of multiple positive blood cultures with the same bacterium. Microbiological tests performed more than 48 hours after ED presentation were excluded.

Demographics, clinical, biochemical and microbiological results were collected from electronic patient records. The Charlson Comorbidity Index was included as a measure for burden of comorbidities.¹⁴ The PSI score and quick sequential organ failure assessment score were used for disease severity at presentation.^{10,15}

Outcomes

The primary outcome was the incidence of bacterial coinfection at ED presentation and its association with PCT. The secondary outcome was severe disease course, defined as admission to the ICU and/or mortality within 30 days. Biomarkers were compared with the PSI score since this is routinely used at our ED to assess the risk of mortality in patients suffering from pneumonia.

Patient and public involvement

No patients were involved in the production of this research.

Statistical methods

Continuous variables are presented as mean $\pm$ SD or median and IQR. Categorical data are summarised as counts and percentages. Baseline characteristics were compared between groups of severity using χ^2 or Fisher's exact test for categorical variables, and the unpaired t-test or Mann-Whitney U test for continuous variables. Univariate and multivariable logistic regression was used to determine the association between all biomarkers and severe outcome, including age as potential confounder. PCT was also corrected for severe kidney disease, since PCT levels are significantly higher in patients with an advanced stage of chronic kidney disease.¹⁶ Severe kidney disease was defined as patients on dialysis, creatinine >270 $\mu\text{mol/L}$, uraemia or a history of kidney transplant. Variables were logarithmically transformed in regression analyses. The log₂ was used; in this way ORs refer to a doubling in

the value of the variable.¹⁷ ORs were also calculated for pre-defined internationally used cut-off values for PCT (<0.10 ng/mL, 0.10–0.25 ng/mL, 0.25–0.50 ng/mL and >0.50 ng/mL) and pre-defined PSI-categories (PSI 1–5). For this analysis, the non-transformed data was used. The area under the receiver operating characteristics (ROC) curve (AUC) was used to quantify discriminative ability for all biomarkers, PSI score, and for a combination of the biomarker with the highest predictive value and the PSI score. Patients with missing values were excluded per analysis. The number of missing values is reported below each table. Statistical significance was assumed for a p value ≤ 0.05 . All analyses were performed using IBM SPSS statistics V.24.0 (IBM Corp).

RESULTS

In total, 485 ED patients had a blood culture taken in March and April 2020, of whom 142 patients tested positive for SARS-CoV-2 (figure 1). The mean age of the 142 patients was 61 years (SD ± 17.1), ranging from 25 to 99 years, of which 66% (93/142) were men. More than half of the patients suffered from any comorbidity, 25% had diabetes, 24% had a history of cardiovascular disease and 20% had a history of pulmonary disease (table 1).

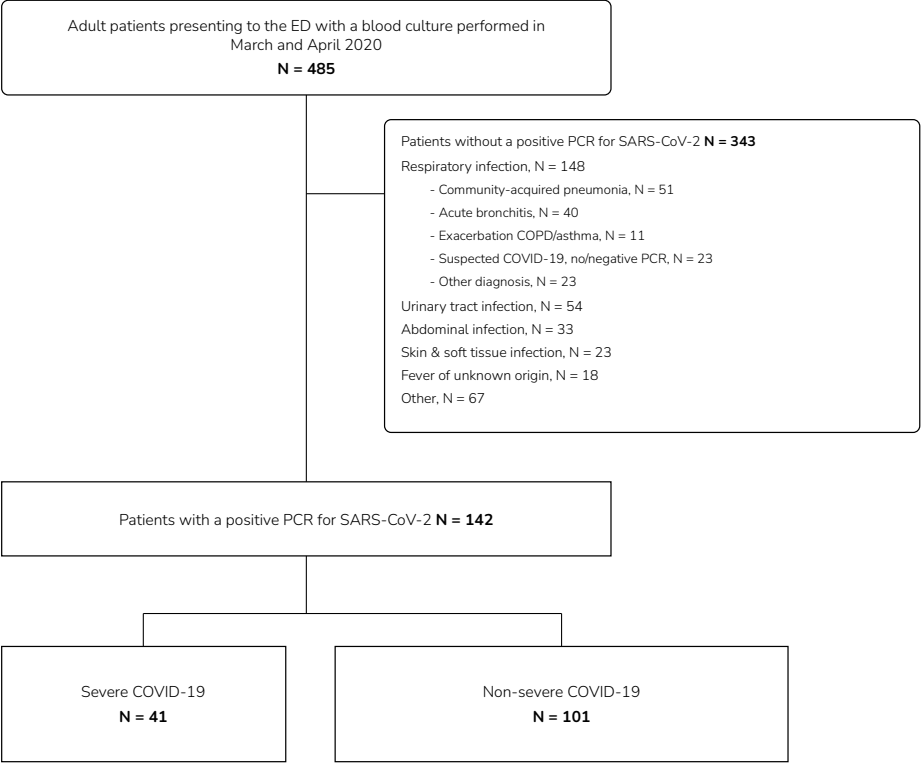


Figure 1 - Flowchart of patient inclusion. ED: emergency department; COPD: chronic obstructive pulmonary disease

Table 1: Baseline characteristics

Demographics	Non-severe, n=101	Severe, n=41	Total, n=142	P-value ^y
Age, <i>mean (SD)</i>	58 (17,1)	69 (14,9)	61 (17,1)	<0.001
Males, <i>no (%)</i>	65 (64,4%)	28 (68,3%)	93 (65,5%)	0.65
Nursing home resident, <i>no (%)</i>	9 (8,9%)	2 (4,9%)	11 (7,7%)	0.51
Comorbidities, <i>no (%)</i>				
Any type of comorbidity	66 (65,3%)	26 (63,4%)	92 (64,8%)	0.83
Diabetes mellitus	22 (21,8%)	13 (31,7%)	35 (24,6%)	0.21
Cardiovascular disease*	23 (22,8%)	11 (26,8%)	34 (23,9%)	0.61
Pulmonary disease [†]	23 (22,8%)	6 (14,6%)	29 (20,4%)	0.28
Malignancy	11 (10,9%)	6 (14,6%)	17 (12%)	0.57
Neurological disease [‡]	10 (9,9%)	6 (14,6%)	16 (11,3%)	0.40
Renal disease [§]	8 (7,9%)	4 (9,8%)	12 (8,5%)	0.74
Charlson Comorbidity Index <i>median (IQR)</i>	2 (0-3,5)	3 (1,5-4,5)	2 (0-4)	<0.01
Symptoms and vital signs at presentation				
Time since onset of symptoms, days, <i>median (IQR)</i>	9 (7-12)	7 (4-11)	8 (6-11)	0.03
Temperature, Celsius, <i>mean (SD)</i>	38,6 (0,9)	38,4 (1)	38,5 (0,9)	0.23
Respiratory rate, breaths/min, <i>mean (SD)</i>	24,9 (6,3)	29,5 (7,6)	26,2 (7)	<0.001
Saturation, %, <i>median (IQR)</i>	95 (93-96)	91 (88-95)	94 (91-96)	<0.001
Heart rate, beats/min, <i>mean (SD)</i>	98 (17)	99 (20)	98 (18)	0.62
Systolic blood pressure, mmHg, <i>mean (SD)</i>	129,9 (21,7)	132,2 (31,9)	130,6 (25)	0.62
Diastolic blood pressure, mmHg, <i>mean (SD)</i>	75,7 (12,7)	75,2 (15,7)	75,6 (13,6)	0.88
Glasgow Coma Score, <i>median (IQR)</i>	15 (15-15)	15 (15-15)	15 (15-15)	0.66
q-SOFA score, <i>mean (SD)</i>	0,7 (0,6)	1 (0,5)	0,8 (0,6)	<0.01
PSI score <i>mean (SD)*</i>	74,2 (31,9)	99,7 (37,5)	82 (35)	<0.001
PSI 1, <i>no (%)</i>	22 (21,8%)	3 (7,3%)	25 (17,6%)	
PSI 2, <i>no (%)</i>	30 (29,7%)	4 (9,8%)	34 (23,9%)	
PSI 3, <i>no (%)</i>	20 (19,8%)	10 (24,4%)	30 (21,1%)	
PSI 4, <i>no (%)</i>	24 (23,8%)	18 (43,9%)	42 (29,6%)	
PSI 5, <i>no (%)</i>	5 (5%)	6 (14,6%)	11 (7,7%)	
Use of antibiotics in last 7 days, <i>no (%)</i>	30 (30%)	9 (22%)	39 (27,7%)	0.33
Use of immunosuppressants, <i>no (%)</i>	5 (5%)	3 (7,3%)	8 (5,6%)	0.69

Laboratory findings at admission, median (IQR)				
Creatinine, $\mu\text{mol/L}$	83 (66-101)	101 (83-137)	87 (69-108)	<0.001
LDH, U/L	284 (224-422)	416 (264-517)	317 (234-441)	<0.01
Ferritin, $\mu\text{g/L}$	534 (263-819)	824 (369-1255)	564 (273-987)	0.03
Trombocytes, $10^9/\text{L}$, <i>mean (SD)</i>	202 (65)	201 (78)	202 (69)	0.94
White blood cell count, $10^9/\text{L}$	6,3 (5-8,1)	6,8 (5-9,6)	6,6 (5-8,3)	0.22
Lymphocyte count, $10^9/\text{L}$	1,02 (0,7-1,3)	0,81 (0,6-1,2)	0,96 (0,7-1,3)	0.03
CRP, mg/L	74 (35-122)	127 (82-203)	86 (44-142)	<0.001
PCT, ng/mL	0,12 (0,08-0,23)	0,3 (0,17-0,78)	0,15 (0,09-0,3)	<0.001
PCT >0.5, ng/mL, <i>no (%)</i>	10 (9,9%)	14 (34,1%)	24 (16,9%)	<0.001
CT value of PCR, <i>mean (SD)</i>	24,6 (4,9)	23,2 (5,4)	24,2 (5,1)	0.17
Outcomes, median (IQR)				
Admission, <i>no (%)</i>	79 (78,2%)	41 (100%)	120 (84,5%)	<0.01
Number of days admitted	3 (2-6)	7 (4-24)	4 (2-8)	<0.001
Start of antibiotics, <i>no (%)</i>	25 (24,8%)	21 (51,2%)	46 (32,4%)	<0.01

* Cardiovascular diseases includes i.a. a history of (un)stable angina pectoris, heart rhythm disorder, myocardial infarction, chronic heart failure, peripheral vascular disease

† Pulmonary disease includes chronic obstructive pulmonary disease, asthma, cystic fibrosis, obstructive sleep apnea syndrome or pulmonary hypertension

‡ Neurological disease includes cerebral vascular attack, transient ischemic attack, multiple sclerosis, Amyotrophic lateral sclerosis and Parkinson disease

§ Renal disease includes an eGFR <50 mL/min, proven kidney disease or kidney transplant

* PSI group 1: 0-50, group 2: 51-70, group 3: 71-90, group 4: 91-130, group 5: 131-395

¥ No adjustment for multiple testing was made

No. of missing values per variable (n): Days since onset of symptoms (16), respiratory rate (7), heart rate (3), systolic and diastolic blood pressure (2), use of antibiotics in last 7 days (1), creatinine (1), LDH (9), ferritin (18), lymphocyte count (6), CT value PCR(8)

Abbreviations: PSI: Pneumonia Severity Index; LDH: lactate dehydrogenase; CRP: C-reactive protein; PCT: procalcitonin; CT: cycle threshold

Microbiological findings

In this cohort, 1.4% (2/142) of the patients had a bacterial coinfection at presentation (table 2); antibiotics were prescribed in 32% (46/142); 28% already received antibiotics prior to presentation to the ED (table 1). In the first 48 hours after presentation, 175 blood cultures were taken, resulting in 11 positive cultures from 10 individuals that were all deemed contaminants (online supplemental appendix A). Because there were only two patients with bacterial coinfection, the association between biomarkers and microbiological outcome could not be further investigated.

Table 2: Microbiological results of all tests performed within 48 hours of presentation

	Total	Positive	Deemed contaminants
Blood cultures	175	11	11
Sputum cultures	15	1	0
<i>Mycoplasma</i> or <i>Chlamydia</i> PCR	2	0	0
Pneumococcal antigen tests	14	1	0
Legionella antigen tests	15	0	0

The first patient had a positive pneumococcal urinary antigen test and a lobar infiltrate on chest X-ray. He was on long-term treatment with prednisone and azathioprine for microscopic polyangiitis. The second patient had a medical history of chronic obstructive pulmonary disease with known colonisation with *Pseudomonas aeruginosa*. In this case, a sputum culture was again positive for *P. aeruginosa*. Neither one had a severe disease course; both recovered after 6 days of admission. A more extensive summary of the clinical presentation of these patients is given in online supplemental appendix B.

Severe disease course

Of 142 patients, 41 patients met the definition for severe disease course; 24 (17%) patients died within 30 days. All patients in the severe disease course group were admitted to the hospital; in the non-severe group, 79 out of 101 (78%) were admitted (table 1). In the severe disease course group, the median Charlson Comorbidity Index was significantly higher compared with the non-severe group (3 vs 2, $p<0.01$).

The median time since onset of symptoms until presentation at our ED was 8 days (IQR 5). At presentation, the patients with a severe disease course had a higher respiratory rate (30 vs 25, $p<0.001$) and lower peripheral oxygen saturation (91% vs 95%, $p<0.001$). There was no significant difference in temperature, blood pressure, heart rate or Glasgow Coma Score. In the laboratory findings, patients with a severe

disease course had significantly higher levels of creatinine, lactate dehydrogenase, ferritin, CRP and PCT, whereas the lymphocyte count was significantly lower.

In logistic regression analysis, CRP and PCT were both significantly associated with severe disease course, with ORs of 1.8 (unadjusted) and 1.7 and 1.8 (adjusted), respectively (table 3). Lymphocyte count was not significantly associated with severe disease course in our study (OR 0.7 (0.4 to 1.1)). For the different PCT categories and the PSI groups, a higher category resulted in higher odds for severe outcome.

Table 3: ORs for biomarkers, PSI and severe outcome

	Odds ratio	95% C.I.	Adjusted OR	Adjusted 95% C.I.
Odds ratio for biomarkers †				
PCT*	1.8	1.3 - 2.3	1.7	1.3 - 2.2
CRP [§]	1.8	1.3 - 2.6	1.8	1.3 - 2.6
Lymphocytes [§]	0.6	0.3 - 0.9	0.7	0.4 - 1.1
Ferritin [§]	1.3	0.99 - 1.8	1.4	1.04 - 1.9
PCT in groups*				
PCT <0.10 ng/mL	Ref	Ref	Ref	Ref
PCT 0.10 - 0.25 ng/mL	6.2	1.7 - 23	5.5	1.5 - 20.7
PCT 0.25-0.50 ng/mL	11.1	2.4 - 51.6	10.5	2.1 - 52.1
PCT >0.50 ng/mL	20.1	4.8 - 83.4	16	3.8 - 68.4
Odds ratio for PSI groups				
PSI group 1	Ref	Ref		
PSI group 2	0.98	0.2 - 4.8		
PSI group 3	3.7	0.9 - 15.2		
PSI group 4	5.5	1.4 - 21.3		
PSI group 5	8.8	1.6 - 47.8		

† The log2 was used, in this way the odds ratios refer to a doubling in the value of the variable
 * Adjusted for age and severe kidney disease (On dialysis, creatinine >270 µmol/L, uraemia, history of kidney transplant)
 § Adjusted for age
 Abbreviations: PCT: procalcitonin; CRP: C-reactive protein; PSI: Pneumonia Severity Index

The ROC curve for PCT had an AUC of 0.76 (0.68 to 0.84), CRP showed an AUC of 0.70 (0.61 to 0.79). The AUC of the PSI score was 0.72 (0.63 to 0.81) (table 4). Combining PCT with PSI did not result in an increase in AUC compared with PSI or PCT alone (0.73, 0.64 to 0.82).

Table 4: AUC of the ROC curve for biomarkers and severe disease course

	AUC	95% C.I.
PCT	0.76	0.68 - 0.84
CRP	0.70	0.61 - 0.79
Ferritin	0.62	0.51 - 0.74
Lymphocyte count	0.62	0.51 - 0.72
PSI score	0.72	0.63 - 0.81

To exclude that our findings were driven by the non-severe subgroup of patients with COVID-19 who were not admitted to the hospital (n=22), the analysis was also performed for hospitalised patients only (n=79 non-severe vs n=41 severe). This revealed similar results (see online supplemental table 2 and table 3).

DISCUSSION

We found that in patients with COVID-19 presenting to the ED from whom a blood culture was taken, bacterial coinfection was only present in 1.4% of the patients, which supports a restrictive antibiotic policy. Moreover, blood cultures seem to be of limited value in this setting as no pathogens were detected whereas 6.3% of the blood cultures returned false positive due to contamination. Of all biomarkers investigated, PCT was most discriminative between severe disease course and non-severe disease course compared with CRP, ferritin and lymphocytes. The widely implemented PSI score also performed well compared with these biomarkers.

A strength of our study is that only patients from whom a blood culture was drawn were included. The low rate of bacterial coinfection is in line with previous studies reporting rates of 0%–8%.^{5,6} However, as blood cultures are indicated based on clinical judgement of the attending physician (eg, suspected sepsis), it is likely that a relatively high risk of bacterial coinfection was suspected. Seen in this light, the 1.4% incidence of bacterial pulmonary coinfection is surprisingly low. However, our study population is different than most previous reports, that mostly refer to non-European, hospitalised patients including also non-respiratory infections.^{5,6} In addition, in the Dutch healthcare system almost all patients consult their general practitioner before visiting the ED. Consequently, 28% of our population received antibiotics prior to ED presentation, which may have decreased the number of detected bacterial coinfections.

The 6.3% rate of contaminated blood cultures in our study is remarkably high. To interpret this number, we analysed contamination rates of all patients included in the PredictED study who were diagnosed with a respiratory infection before the COVID-19 pandemic. In this period, 2.6% (16/604) of blood cultures were

contaminated (data not shown). One could argue that the high work pressure at the ED during the first wave of COVID-19 and the burden of working in protective equipment might have contributed to a higher rate of contamination.

Recently, Karami et al reported a bacterial coinfection rate of 1.2% in the first week of hospitalisation in 925 patients. In this study, only four blood cultures were positive while 60 from 711 blood cultures (8.4%) were contaminated.¹⁸ Another retrospective study in the Netherlands among patients with COVID-19 with a blood culture being performed within 48 hours after a positive SARS-CoV-2 PCR test showed a bacteraemia rate of 1% while 8% of blood cultures were contaminated.¹⁹ In a large multicentre study conducted in New York City, the yield of blood cultures in patients with COVID-19 was only 1.6% while another 2.2% was contaminated.²⁰ This included pathogens indicating non-pulmonary coinfections and blood cultures that were taken either at the ED or during hospitalisation. As in our study, the number of false positive blood cultures exceeded the number of true positives in these studies. Meanwhile, the large numbers of blood cultures overwhelmed the capacity of laboratories. In addition, contamination of blood cultures is associated with a prolonged hospital stay and more antibiotic prescription.²¹ Therefore, we suggest that blood cultures should not be routinely performed in patients with COVID-19 presenting to the ED. Sputum cultures and urinary antigen tests might be a more efficient way to diagnose bacterial coinfection, especially in patients with known underlying pulmonary or immunocompromising comorbidities, as in our study. However, with a yield of 1.4%, our study does not support the routine use of sputum cultures and antigen testing for all patients with COVID-19.

PCT is known as a marker of bacterial infection.²² In our cohort, PCT levels ranged from 0.02 to 25 ng/mL. Since we found bacterial coinfection in only two patients, this seems to play a limited role in the elevated PCT levels of patients with COVID-19 presenting to the ED. Likely, in absence of bacterial infection, this might reflect systemic inflammation, which is also known to increase PCT levels.²³ This would be consistent with the beneficial effects of corticosteroids in patients with COVID-19 in need of oxygen therapy, which is aimed at tempering host immune response.²⁴ Because of the limited number of bacterial coinfections, we were not able to investigate the association with PCT value in this population. However, our results suggest that even in patients with elevated PCT levels or a high PSI score, a restrictive antibiotic policy could still be considered. Whether PCT can be used to rule out a bacterial coinfection should be investigated in future studies.

In previous studies, the biomarkers PCT, CRP, ferritin and lymphocyte count were found to be associated with a severe disease course.^{3,4} However, to our knowledge, the discriminative ability of these biomarkers has not been compared. Many studies

have tried to develop a prediction model for disease course severity in COVID-19. Unfortunately, as illustrated in a systematic review by Wynants et al, almost all models were at high risk of bias and thus are not recommended to use in current practice.²⁵ Only the 4C Mortality Score, which includes CRP as a biomarker, is considered to be promising.²⁶ Our results may provide guidance when assumptions must be made in developing a prediction model. For instance, it would be of interest to evaluate the added value of PCT in the 4C Mortality Score.

Our study has several limitations. First, it is an observational study in which the treating physician decided whether to perform microbiological tests. Ideally, the rate of bacterial coinfection would be determined by performing a standardised set of tests in all patients. However, bacterial pulmonary coinfection is diagnosed by combining clinical signs and test results, so including the physician's judgement seems reasonable. Second, our definition of bacterial coinfection is based on microbiological test results only, ignoring a potential clinical diagnosis or a negative culture because of antibiotic pre-treatment. Nevertheless, our study underlines the limited value of blood cultures in diagnosing pulmonary coinfection.^{7,27} Third, our sample size is relatively small, so our findings should be interpreted with caution. However, our findings are in accordance with other studies with larger sample sizes.^{6,18,19} Finally, only patients from whom a blood culture was drawn were included, so we were not able to study patients with COVID-19 who had no blood culture taken at ED. However, it is likely this would only decrease the rate of bacterial coinfection as there was no clinical suspicion.

In conclusion, the rate of bacterial coinfection of patients with COVID-19 presenting to the ED is low, which supports a restrictive antibiotic policy. Blood cultures appear to have limited value while PCT and the PSI appear promising tools in helping physicians recognise patients at risk for severe disease course.

Data availability statement

The deidentified participant data that support the findings of this study are available from the corresponding author, upon reasonable request. Email address: a.g.kaal@hagaziekenhuis.nl.

Ethics statements

Patient consent for publication

Not required.

Ethics approval

The PredictED study was approved by the Medical Ethics Committee and the Institutional Scientific Review Board (T18-040). Informed consent was waived since it is an observational quality control study without impact upon patients' medical management.

Contributors

AK, LS, SM and CvN contributed to the design of this research. Together with ESté they were responsible for the analysis. AK and LS wrote the draft of the article. CvN directed and supervised the project. All authors provided critical feedback, reviewed and revised the manuscript.

Funding

This study was partly sponsored by an unrestricted grant (T18-040) of the Haga Scientific Research Fund.

Competing interests

TO reports personal lecture fees from Gilead, outside the submitted work.

SUPPLEMENTAL MATERIAL

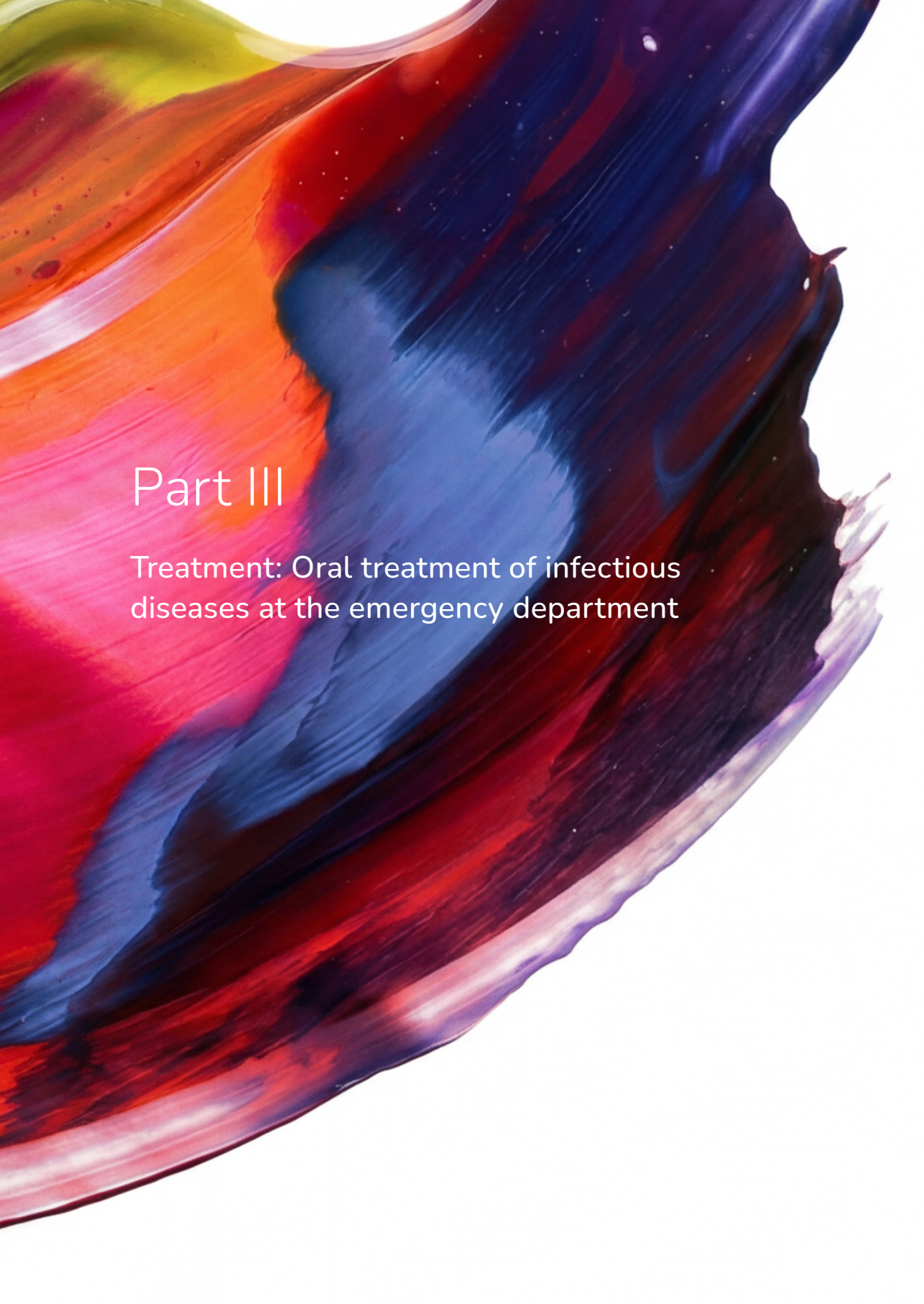


REFERENCES

1. Guan, Ni, Yu Hu, W. Liang, Ou, He, L. Liu, Shan, Lei, Hui, Du, L. Li, Zeng et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020;382:1708-20. DOI: 10.1056/NEJMoa2002032
2. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3
3. Lippi G, Plebani M. Procalcitonin in patients with severe coronavirus disease 2019 (COVID-19): A meta-analysis. *Clin Chim Acta*. 2020;505(March):190-191. doi:10.1016/j.cca.2020.03.004
4. Malik P, Patel U, Mehta D, et al. Biomarkers and outcomes of COVID-19 hospitalisations. Systematic review and meta-analysis. *BMJ Evid Based Med* 2021;26:107-8
5. Rawson TM, Moore LSP, Zhu N, et al. Bacterial and fungal co-infection in individuals with coronavirus: a rapid review to support COVID-19 antimicrobial prescribing. *Clin Infect Dis*. Published online May 2, 2020. doi:10.1093/cid/ciaa530
6. Langford BJ, So M, Raybardhan S, et al. Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis. *Clin Microbiol Infect*. 2020.
7. Müller et al. Procalcitonin Levels Predict Bacteremia in Patients With Community-Acquired Pneumonia A Prospective Cohort Trial. *CHEST* 2010; 138(1):121-129
8. Pfister et al. Procalcitonin for diagnosis of bacterial pneumonia in critically ill patients during 2009 H1N1 influenza pandemic: a prospective cohort study, systematic review and individual patient data meta-analysis. *Critical Care* 2014, 18:R44 <http://ccforum.com/content/18/2/R44>
9. Metlay JP, Waterer GW, Long AC, et al. Diagnosis and treatment of adults with community-acquired pneumonia. *Am J Respir Crit Care Med*. 2019;200(7):E45-E67. doi:10.1164/rccm.201908-1581ST
10. Fine et al. A Prediction Rule to Identify Low-Risk Patients With Community-Acquired Pneumonia. *N Engl J Med* 1997;336:243-50.
11. Wiersinga et al. Management of Community-Acquired Pneumonia in Adults: 2016 Guideline Update From The Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). *Neth J Med*. 2018 Jan;76(1):4-13.
12. Sieswerda E et al., Recommendations for antibacterial therapy in adults with COVID-19, an evidence based guideline, *Clinical Microbiology and Infection*, <https://doi.org/10.1016/j.cmi.2020.09.041>
13. Corman Victor M, Landt Olfert, Kaiser Marco, Molenkamp Richard, Meijer Adam, Chu Daniel KW, Bleicker Tobias, Brünink Sebastian, Schneider Julia, Schmidt Marie Luisa, Mulders Daphne GJC, Haagmans Bart L, van der Veer Bas, van den Brink Sharon, Wijsman Lisa, Goderski Gabriel, Romette Jean-Louis, Ellis Joanna, Zambon Maria, Peiris Malik, Goossens Herman, Reusken Chantal, Koopmans Marion PG, Drosten Christian. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*. 2020;25(3):pii=2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>
14. M.E. Charlson, P. Pompei, K.L. Ales, C.R. MacKenzie A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis*, 40 (1987), pp. 373-383, 10.1016/0021-9681(87)90171-8

15. Singer et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) JAMA. 2016 February 23; 315(8): 801–810. doi:10.1001/jama.2016.0287.
16. Grace E, Turner R. Use of Procalcitonin in Patients With Various Degrees of Chronic Kidney Disease Including Renal Replacement Therapy. *Clinical Infectious Diseases* 2014;59(12):1761–7
17. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. Springer; 2019.
18. Karami et al. (2021) Few bacterial co-infections but frequent empiric antibiotic use in the early phase of hospitalized patients with COVID-19: results from a multicentre retrospective cohort study in The Netherlands, *Infectious Diseases*, 53:2, 102-110, DOI: 10.1080/23744235.2020.1839672
19. Thelen et al. (2021) Community-acquired bacteraemia in COVID-19 in comparison to influenza A and influenza B: a retrospective cohort study. *BMC Infectious Diseases* (2021) 21:199. <https://doi.org/10.1186/s12879-021-05902-5>
20. Sepulveda et al. Bacteremia and Blood Culture Utilization During COVID-19 Surge in New York City. *J. Clin. Microbiol.* doi:10.1128/JCM.00875-20
21. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization. The true consequences of false-positive results. *JAMA*. 1991 Jan 16;265(3):365-9. PMID: 1984535.
22. Schuetz et al. Procalcitonin for diagnosis of infection and guide to antibiotic decisions: past, present and future. *BMC Medicine* 2011, 9:107 <http://www.biomedcentral.com/1741-7015/9/107>
23. Kenneth L. Becker, MD, PhD; Richard Snider, PhD; Eric S. Nylen, MD. Procalcitonin assay in systemic inflammation, infection, and sepsis: Clinical utility and limitations. *Crit Care Med* 2008 Vol. 36, No. 3
24. RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with Covid-19—Preliminary report. *N Engl J Med NEJMoa2021436*, 2020. doi:10.1056/NEJMoa2021436.
25. Wynants et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328 <http://dx.doi.org/10.1136/bmj.m1328>
26. Knight et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ*2020;370:m3339
27. Samuel G. Campbell, Thomas J. Marrie, Rosemary Anstey, Garth Dickinson, Stacy Ackroyd-Stolarz, The Contribution of Blood Cultures to the Clinical Management of Adult Patients Admitted to the Hospital With Community-Acquired Pneumonia: A Prospective Observational Study, *Chest*, Volume 123, Issue 4, 2003, Pages 1142-1150, ISSN 0012-3692, <https://doi.org/10.1378/chest.123.4.1142>.





Part III

Treatment: Oral treatment of infectious diseases at the emergency department



Chapter 8

Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study

Anna G. Kaal*, Rick Roos*, Pieter de Jong, Rianne M. C. Pepping, Johanna M. W. van den Berg, Maarten O. van Aken, Ewout W. Steyerberg, Mattijs E. Numans, Cees van Nieuwkoop

*Both authors contributed equally

ABSTRACT

Community-acquired Pneumonia (CAP) guidelines generally recommend to admit patients with moderate-to-severe CAP and start treatment with intravenous antibiotics. This study aims to explore the clinical outcomes of oral antibiotics in patients with moderate-to-severe CAP. We performed a nested cohort study of an observational study including all adult patients presenting to the emergency department of the Haga Teaching Hospital, the Netherlands, between April 2019 and May 2020, who had a blood culture drawn. We conducted propensity score matching with logistic and linear regression analysis to compare patients with moderate-to-severe CAP (Pneumonia Severity Index class III–V) treated with oral antibiotics to patients treated with intravenous antibiotics. Outcomes were 30-day mortality, intensive care unit admission, readmission, length of stay (LOS) and length of antibiotic treatment. Of the original 314 patients, 71 orally treated patients were matched with 102 intravenously treated patients. The mean age was 73 years and 58% were male. We found no significant differences in outcomes between the oral and intravenous group, except for an increased LOS of + 2.6 days (95% confidence interval 1.2–4.0, p value < 0.001) in those treated intravenously. We conclude that oral antibiotics might be a safe and effective treatment for moderate-to-severe CAP for selected patients based on the clinical judgement of the attending physician.

INTRODUCTION

Community-acquired pneumonia (CAP) is one of the main causes of mortality and hospitalisation and is usually treated with one or more types of antibiotics.¹⁻³ CAP has an estimated admission rate of 50–2.940 per 100,000 annually in the adult population of Europe, and is highly dependent of age.^{4,5} The costs of CAP are estimated to be around 10.1 billion euros annually in Europe alone, most of which comes from inpatient care.³

In current European and American guidelines, treatment location (inpatient or outpatient) and route of antibiotic administration (oral or intravenous) are based on a combination of clinical prediction rules and clinical judgement of the physician.⁶⁻¹⁰ These prediction rules (e.g. the Pneumonia Severity Index (PSI) and CURB-65) assess disease severity and divide patients in different categories according to their mortality risk.^{11,12} For moderate-to-severe CAP (PSI class III–V, or CURB-65 score ≥ 2), most guidelines advise hospitalisation and treatment with intravenous antibiotics.⁶⁻¹⁰

Some studies have explored the possibility of treating CAP patients with oral antibiotics.¹³⁻¹⁵ However, most of these studies did not specify the severity of disease or only included highly selected patient groups with mild-to-moderate CAP. Two studies included hospitalised patients with moderate-to-severe CAP and showed similar or even better outcomes for oral treatment compared to intravenous treatment. Both studies did not specify disease severity in patients with treatment failure.^{14,15} Furthermore, one study showed that a substantial number of patients presenting to the emergency department (ED) with CAP PSI class IV–V can be treated at home safely.¹⁶

In this study, we aim to explore the clinical outcomes of oral versus intravenous antibiotic treatment in patients with moderate-to-severe CAP presenting to the ED in an observational cohort of patients representing current daily practice in a large teaching hospital in the Netherlands.

MATERIALS AND METHODS

Study design and setting

This is a nested cohort study of a single-centre observational study, the PredictED study. The PredictED study aims to predict bacteraemia in patients presenting to the ED with suspected infection. All patients aged 18 years or older from whom a blood culture was taken during an ED-visit between April 2019 and May 2020 were included. This study was conducted at the ED of the Haga Teaching Hospital in the

Netherlands, a large urban hospital with around 600 beds and 50.000 ED-visits annually. The PredictED has been registered at the national trial register: [<https://onderzoekmetmensen.nl/nl/trial/23345>].

We included patients with an ED diagnosis of moderate-to-severe CAP, defined as PSI class III–V. If patients visited the hospital multiple times, we only included the first visit. We only included patients treated with antibiotics for pneumonia within 24 h of the ED-visit. Furthermore, we excluded patients diagnosed with a COVID-19 infection, an aspiration pneumonia or a hospital-acquired pneumonia (HAP), since these diagnoses require a different treatment approach than for CAP.¹⁰ Aspiration pneumonia was defined as every pneumonia in which, as according to the electronic health record (EHR), the attending physician strongly considered aspiration to be the underlying cause and subsequent empirical treatment was started (according to the local hospital guideline being amoxicillin clavulanic acid or ceftriaxone combined with metronidazole). Furthermore, patients who were incurably ill due to other diseases (e.g. cancer) and opted for palliative care after their ED visit were excluded. Finally, patients who were transferred to other hospitals were excluded.

Treatment

Treatment of CAP was left to the attending physician at the ED. The physicians followed the local hospital guideline, and could deviate from the first choice in the guideline based on their clinical judgement (e.g. in case of known pathogens from previous positive sputum cultures or suspicion of an atypical respiratory pathogen). Our local hospital guideline recommends treating PSI class III–IV pneumonia with oral amoxicillin 750 mg, 3 times a day (total treatment duration 5 days); or intravenous amoxicillin 1000 mg, 4 times a day (total treatment duration 5 days). The choice for oral or intravenous treatment is based on the clinical judgement of the attending physician at the ED, the local guideline does not contain criteria for clinical judgement. In case of pre-treatment with amoxicillin or failure of amoxicillin treatment, the local guideline recommends an oral switch to doxycycline 100 mg, once a day with a loading dose of 200 mg on the first day (total treatment duration 7 days) or moxifloxacin 400 mg, once a day (total treatment duration 5 days). The first choice of treatment of patients with PSI class V pneumonia is ceftriaxone 2000 mg, once a day (total treatment duration 5 days). In case of clinical suspicion of an atypical respiratory pathogen (e.g. *Legionella* spp.), the addition of a fluoroquinolone is recommended. Based on the pharmacokinetic profiles, our local antibiotic hospital guideline committee considers there to be no rationale for treatment with intravenous amoxicillin or moxifloxacin over oral amoxicillin or moxifloxacin.^{17–22} Therefore, our local hospital guideline differs from the Dutch national guideline in the treatment of

patients with PSI class III–IV pneumonia as the latter recommends similar antibiotics albeit always intravenously.¹⁰ This difference with the national guideline gave us the unique opportunity to compare oral and intravenous treatment for this patient category.

Variables and data collection

Demographical, clinical, biochemical, and microbiological data were collected from the EHR. This included the resuscitation code status, describing the type of resuscitation procedures (cardiopulmonary resuscitation, mechanical ventilation, and ICU-admission) used in case of a medical emergency. In case a code status was not documented in the EHR at the time of the ED-visit, two physicians (AK, RR) reviewed the EHRs to evaluate whether there would have been possible reasons to restrict the code status to a non-ICU-admission code.

We defined cerebrovascular disease, congestive heart failure, liver disease, malignancy and renal disease according to the definitions of the Charlson Comorbidity Index.²³ In addition, pulmonary disease was defined as the presence of chronic pulmonary disease including asthma, chronic obstructive pulmonary disease, cystic fibrosis, pulmonary hypertension, or obstructive sleep apnoea. Besides that, immunosuppressant use was defined as the use of medication suppressing the immune system, including the chronic use of glucocorticosteroids, cyclosporine, methotrexate, azathioprine, cyclophosphamide, tacrolimus and biologicals. Recent antibiotic use was defined as use of antibiotics up to seven days prior to ED-visit. Finally, we defined altered mental status as a Glasgow Coma Scale < 15 points or a note of altered mental status.²⁴

We collected data on antibiotic treatment regimens. Patients were classified as either 'Oral' or 'Intravenous' according to the method of administration of the first antibiotic regimen. Patients who were given only one dose of intravenous antibiotics and then switched to an oral antibiotic were categorised in the oral treatment group as it is not uncommon at our ED to administer the first dose of antibiotics intravenously before starting oral treatment. Therapeutic antibiotic switches were defined as a change of antibiotic treatment to a more broad-spectrum or entirely different spectrum of antibiotics, or a change from oral to intravenous antibiotics.

Biochemical data consisted of the laboratory test results at the ED required to calculate the PSI-score and inflammatory markers (white blood cell count and C-reactive protein (CRP)). In addition, the presence of pleural effusion on the chest X-ray was documented. Based on these demographical, clinical, and biochemical variables the PSI-score was calculated for all patients.

Microbiological data consisted of all blood and sputum cultures, polymerase chain reactions (PCR) of nasopharyngeal swabs for viral respiratory pathogens, urinary antigen tests and atypical respiratory pathogen tests taken within 48 h of the ED-visit.

Outcomes

The primary study outcome was 30-day mortality. Secondary study outcomes were ICU-admissions, unplanned readmissions within seven days of discharge, length of hospital stay (LOS) and length of antibiotic treatment.

Statistical analysis

Descriptive statistics were presented as numbers and percentages, means and standard deviations (SD) or medians and interquartile ranges (IQR), as appropriate. A two-sided p value < 0.05 was considered statistically significant.

Missing data was imputed ten times using Multiple Imputation by Chained Equation (MICE).²⁵ Rubin's rules were then used to create a single dataset.²⁶ Since blood pH is most likely not to be missing at random, we assumed missing values meant a blood pH > 7.35 , as it was the clinical judgement of the physician to not perform a blood gas analysis. Since the assignment of oral or intravenous treatment was not random, we performed a propensity score weighted analysis. We estimated the propensity score with a binary logistic regression model consisting of all individual parameters as used in the PSI, using the absolute values where possible, history of pulmonary disease, use of immunosuppressants, recent use of antibiotics, non-ICU code, CRP, peripheral oxygen saturation and the need for oxygen therapy. Matching was done using the nearest neighbour method without replacement, with a 1:2 ratio and a caliper of 0.2 of the SD of the logit of the propensity score.²⁷ This created a population in which potential confounders were balanced, defined as a standardised mean difference less than 10%.²⁷ Finally, for the analysis of primary and secondary outcomes we performed Firth's logistic regression analysis and linear regression analysis on the matched cohort, with correction of the propensity score using a spline function. We performed the LOS analysis for all patients and for only admitted patients.

We conducted multiple post-hoc sensitivity analyses: one among all admitted patients whose diagnosis at discharge remained a respiratory infection and were therefore not misdiagnosed.²⁸ Other sensitivity analysis consisted of only admitted patients, only non-deceased patients and an analysis using only the patients in the oral group that did not receive a first, intravenous dose of antibiotics at the

ED. Finally, we performed regression analysis using traditional correction for confounders, using the same variables as used for the propensity score model.

We reported descriptive outcomes according to multiple subgroups; per PSI class, per bacteremia status, per sepsis status (defined as a qSOFA of 2 or more), per viral coinfection status and according to a level of procalcitonin ≥ 0.25 ng/mL.²⁹ We reported microbiological outcomes and antibiotic treatments for the matched cohort and the complete cohort. Statistical analyses were performed using R Statistical Software version v4.2.2; R Core Team 2022.

Ethics

The PredictED study was approved by the Medical Ethics Committee and the Institutional Scientific Review Board of the Haga Teaching Hospital (protocol T18-040). The need for informed consent was waived by The Medical Ethics Committee Leiden The Hague Delft. All methods were performed in accordance with relevant guidelines and regulations. In this study, we adhered to the STROBE reporting guideline for cohort studies.

RESULTS

Of the 2332 patients presenting to the ED, 314 were diagnosed with a moderate-to-severe CAP. Of those, 88 were treated orally and 226 intravenously (see Fig. 1). After propensity score matching (PSM), 71 patients in the oral group were matched with 102 patients in the intravenous group.

The oral and intravenous group were similar regarding demographics, comorbidities, and code status (Table 1). In the original cohort, more patients in the intravenous group presented with an altered mental status, a higher heart frequency and respiratory rate, and a lower systolic blood pressure and oxygen saturation than the oral group. The median PSI score was higher in the intravenous group (112 vs. 96). 33 out of 76 patients with PSI class III were treated orally (43%) while most patients with PSI class V were treated intravenously (62/67, 93%). The entire intravenous group was hospitalised, while 69% of the patients in the oral group were admitted. After PSM, all variables were balanced across the treatment groups except for systolic blood pressure and recent use of antibiotics (See love plot, Appendix 1).

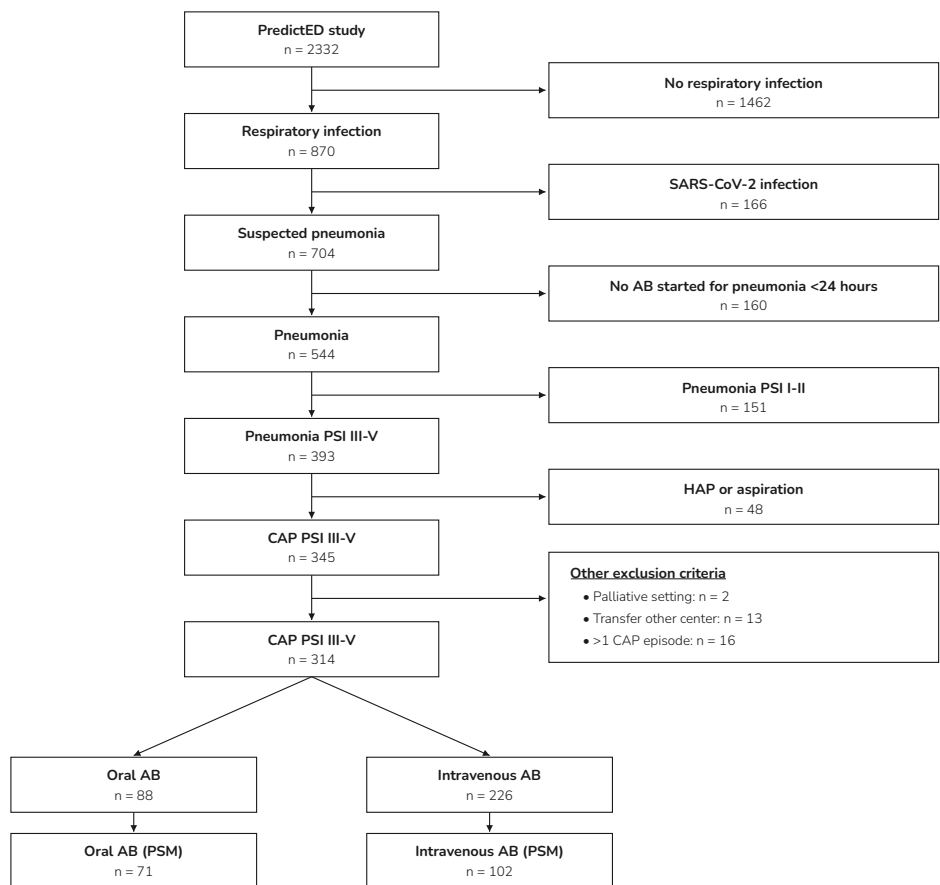


Figure 1: Flowchart of inclusion. *AB* antibiotics, *CAP* community-acquired pneumonia, *HAP* hospital-acquired pneumonia, *PSI* pneumonia severity index, *PSM*: propensity score matching.

Primary outcome

In total, 37/314 patients (11.8%) died during this study, 33 of whom were treated intravenously (Table 1, Appendix 2). In the propensity score matched population, 4/71 patients (5.6%) died in the oral group and 8/102 patients (7.8%) in the intravenous group. Three out of four patients in the oral group died due to causes related to the pneumonia, while the other patient died due to progression of lung cancer. All eight patients in the intravenous group died due to causes related to the pneumonia. We did not find a difference in mortality for intravenous versus oral treatment (aOR 1.1 (0.3–4.7), $p = 0.84$, Table 2).

Secondary outcomes

A total of 29/314 patients (9.2%) were admitted to the ICU and 22 patients (7.0%) were readmitted to the hospital within seven days (Table 1, Appendix 2). No significant differences were seen in ICU-admission rate (aOR 2.8 (0.6–73.2), $p = 0.23$) and readmission rate (aOR 1.7 (0.5–6.8), $p = 0.37$) between both PSM groups (Table 2).

In the matched study population, LOS was significantly longer (4 days vs. 2 days) in the intravenous group (adjusted coefficient: 2.6 days (1.2–4.0), $p < 0.001$). When limiting the analysis to hospitalised patients, no significant difference in LOS was found (adjusted coefficient: 1.4 days (– 0.3 to 3.0), $p = 0.10$). Finally, length of antibiotic treatment did not differ significantly between the oral and intravenous group (7.0 days vs. 6.5 days, adjusted coefficient: 2.2 days (– 0.1 to 4.6), $p = 0.06$).

Mortality increased with every PSI class for both groups (Table 3, Appendix 2). In the matched intravenous group, 6/88 (6.8%) patients with a PSI class III/IV were admitted to the ICU compared to 1/66 (1.5%) in the matched oral group. No orally treated patients with a PSI V were readmitted to our hospital.

Table 1 Baseline characteristics.

	Unmatched		Matched		aSMD	Intravenous (n = 102)
	Oral (n = 88)	Intravenous (n = 226)	Oral (n = 71)	Intravenous (n = 102)		
Demographics						
Age in years, mean (SD)	72 (11)	73 (12)	73 (11)	73 (12)	0.022	0.022
Male, n (%)	50 (56.8)	131 (58.0)	41 (57.7)	60 (58.8)	0.023	0.028
No ICU-admission code, n (%)	31 (35.2)	92 (40.7)	28 (39.4)	36 (35.3)	0.112	0.089
Nursing home resident, n (%)	8 (9.1)	32 (14.2)	8 (11.3)	14 (13.7)	0.145	0.074
Comorbidities						
Cerebrovascular disease, n (%)	15 (17.0)	48 (21.2)	14 (19.7)	19 (18.6)	0.090	0.019
Congestive heart failure, n (%)	17 (19.3)	40 (17.7)	12 (16.9)	19 (18.6)	0.042	0.054
Liver disease, n (%)	2 (2.3)	3 (1.3)	1 (1.4)	2 (2.0)	0.083	0.095
Malignancy, n (%)	25 (28.4)	60 (26.5)	21 (29.6)	27 (26.5)	0.042	0.063
Pulmonary disease, n (%)	39 (44.3)	99 (43.8)	32 (45.1)	46 (45.1)	0.010	0.028
Renal disease, n (%)	9 (10.2)	31 (13.7)	9 (12.7)	12 (11.8)	0.101	0.093
CCI, median [IQR]	4 [3-6]	5 [3-6]	5 [3-6]	4 [3-6]	0.043	0.060
Immunosuppressant use, n (%)	17 (19.3)	38 (16.8)	14 (19.7)	19 (18.6)	0.067	0.018
Recent antibiotic use, n (%)	29 (33.0)	60 (26.5)	25 (35.2)	26 (25.5)	0.145	0.150
Physical examination						
Altered mental status, n (%)	3 (3.4)	47 (20.8)	3 (4.2)	5 (4.9)	0.428	0.039
HF (beats/minute), mean (SD)	102 (18)	110 (24)	102 (19)	106 (22)	0.306	0.090
Oxygen saturation (%), median [IQR]	94 [91-96]	93 [90-95]	94 [91-96]	94 [92-96]	0.293	0.038
Oxygen therapy, n (%)	11 (12.5)	75 (33.2)	11 (15.5)	18 (17.6)	0.439	0.064
RR (breaths/minute), median [IQR]	22 [18-26]	26 [20-30]	22 [18-26]	24 [18-30]	0.368	0.032
SBP (mmHg), mean (SD)	138 (29)	129 (30)	141 (29)	137 (28)	0.317	0.166

Temperature (°C), mean (SD)	38.4 (0.7)	38.6 (1.1)	0.159	38.5 (0.6)	38.6 (1.0)	0.081
Additional diagnostics						
Blood pH, median [IQR]	7.46 [7.45-7.47]	7.46 [7.43-7.48]	0.184	7.46 [7.44-7.46]	7.46 [7.44-7.47]	0.052
CRP (mg/L), median [IQR]	67 [34-170]	125 [47-268]	0.370	76 [38-178]	109 [44-196]	0.030
Glucose (mmol/L), median [IQR]	7.1 [6.2-8.7]	7.6 [6.6-9.4]	0.233	7.3 [6.4-8.9]	7.3 [6.2-9.0]	0.011
Hematocrit < 30%, n (%)	6 (6.8)	20 (8.8)	0.072	4 (5.6)	7 (6.9)	0.000
Sodium (mmol/L), median [IQR]	137 [135-138]	135 [132-138]	0.316	136 [134-138]	136 [133-138]	0.068
Urea (mmol/L), median [IQR]	6.1 [4.9-7.7]	7.8 [5.8-11.7]	0.447	6.6 [5.1-8.7]	6.2 [4.9-8.8]	0.036
WBCs (n x 10 ⁹ /L), median [IQR]	11.8 [8.7-15.2]	12.8 [8.8-17.8]	0.280	11.9 [9.4-15.4]	12.7 [8.9-16.6]	0.183
pO ₂ < 60 mmHg or SaO ₂ < 90%, n (%)	20 (22.7)	58 (25.7)	0.067	14 (19.7)	21 (20.6)	0.034
Pleural effusion on X-ray, n (%)	16 (18.2)	48 (21.2)	0.075	13 (18.3)	19 (18.6)	0.000
Pneumonia severity index (PSI)						
<i>PSI class</i>						
III, n (%)	33 (37.5)	43 (19.0)		22 (31.0)	26 (25.5)	
IV, n (%)	50 (56.8)	121 (53.5)		44 (62.0)	62 (60.8)	
V, n (%)	5 (5.7)	62 (27.4)		5 (7.0)	14 (13.7)	
PSI score, median [IQR]	96 [85-109]	112 [94-133]	0.748	97 [87-110]	103 [90-118]	0.219
Outcomes						
30-day mortality, n(%)	4 (4.5)	33 (14.6)		4 (5.6)	8 (7.8)	
Admission, n (%)	61 (69.3)	226 (100.0)		49 (69.0)	102 (100.0)	
ICU-admission, n (%)	1 (1.1)	28 (12.4)		1 (1.4)	6 (5.9)	
Readmission, n (%)	5 (5.7)	17 (7.5)		4 (5.6)	9 (8.8)	

Table 1: Baseline characteristics before and after propensity score matching. Variables with an aSMD lower than 0.10 are considered to be balanced. aSMD: Absolute standardised mean difference; CCI: Charlson Comorbidity Index. CRP: C-reactive protein; HF: Heart frequency; ICU: Intensive care unit; IQR: Interquartile range; RR: Respiratory rate; SBP: Systolic blood pressure; SD: Standard deviation; WBC: White blood cells. The number of missing values for continuous variables in the unmatched oral group were RR (n = 6), SBP (n = 2), glucose (n = 1) and blood pH (n = 31), and for the unmatched intravenous group these were: RR (n = 7), glucose (n = 3), sodium (n = 1), urea (n = 3) and blood pH (n = 57).

Table 2 Clinical outcomes of oral versus intravenous treatment for patients with CAP.

A									
	Oral (n = 71)	Intravenous (n = 102)	OR	95% CI	p-value	aOR	95% CI	p-value	
Mortality, n (%)	4 (5.6)	8 (7.8)	1.2	0.4-5.0	0.76	1.1	0.3-4.7	0.84	
ICU-admission, n (%)	1 (1.4)	6 (5.9)	3.0	0.6-80.1	0.23	2.8	0.6-73.2	0.23	
Readmission within 7 days, n (%)	4 (4.6)	9 (8.8)	1.7	0.6-6.9	0.36	1.7	0.5-6.8	0.37	
B									
	Oral (n = 71)	Intravenous (n = 102)	β	95% CI	p-value	a β	95% CI	p-value	
Length of stay (days), median [IQR] ^a	2 [0-4]	4 [2-6]	2.6	1.2-4.0	<0.001	2.6	1.2-4.0	<0.001	
Length of stay (days), median [IQR] ^b	3 [2-6]	4 [2-6]	1.3	-0.3-3.0	0.12	1.4	-0.3-3.0	0.10	
Length of antibiotic treatment (days), median [IQR]	7 [5-7.5]	6.5 [5.3-9]	2.1	-0.2-4.4	0.08	2.2	-0.1-4.6	0.06	

Table 2: (A) Logistic regression (B) Linear regression. CI: Confidence interval; ICU: Intensive care unit; OR: Odds ratio, β = effect size in days. aOR and a β : Odds ratio and effect size adjusted for the propensity score. aOR: Adjusted odds ratio; CI: Confidence interval; ICU: Intensive care unit; OR: Odds ratio ^a Outcomes for only non-deceased patients: 67 treated orally, 93 treated intravenously. ^b Outcomes for only hospitalised, non-deceased patients: 45 patients treated orally, 93 treated intravenously.

Table 3 Primary outcomes per PSI-class.

	Oral (n = 71)	Intravenous (n = 102)
Mortality, n (%)		
PSI III	0/22 (0.0)	0/26 (0.0)
PSI IV	3/44 (6.8)	6/62 (9.7)
PSI V	1/5 (20.0)	2/14 (14.3)
ICU-admission, n (%)		
PSI III	0/22 (0.0)	1/26 (3.8)
PSI IV	1/44 (2.3)	5/62 (8.1)
PSI V	0/5 (0.0)	0/14 (0.0)
Readmission within 7 days, n (%)		
PSI III	1/22 (4.5)	2/26 (7.7)
PSI IV	3/44 (6.8)	7/62 (11.3)
PSI V	0/5 (0.0)	0/14 (0.0)
Length of stay (days), median [IQR]		
PSI III	1 [0-2.8]	3 [2-4.8]
PSI IV	2 [0-5]	4 [3-7]
PSI V	2.5 [0.8-5]	4.5 [2.8-7]
Length of stay (days) ^a , median [IQR]		
PSI III	2 [1-4]	3 [2-4.8]
PSI IV	3 [2-6]	4 [3-7]
PSI V	4 [2.5-6]	4.5 [2.8-7]
Length of antibiotic treatment (days), median [IQR]		
PSI III	7 [5-7]	6.5 [6-7.8]
PSI IV	7 [5-8]	7 [5-9]
PSI V	7 [5-7]	6 [5.3-8]

Table 3: ICU: Intensive care unit; IQR: Interquartile range; PSI: Pneumonia severity index. ^a Only hospitalised patients.

All post-hoc sensitivity analyses showed similar results for all primary and secondary outcomes (Appendices 3–7). Descriptive outcomes according to the predefined subgroups can be found in Appendices 8–11.

Microbiology

In the matched study population, blood cultures were positive in 3/71 patients in the oral group (4.2%) and 12/102 patients in the intravenous group (11.8%). In blood

cultures, the most common respiratory pathogen was *Streptococcus pneumoniae*. Sputum cultures were performed in 29/71 orally treated patients (41%) and 44/102 intravenously treated patients (43%). In the sputum cultures, the most common respiratory pathogen in both groups was *Haemophilus influenzae*. Viral swabs were performed in 40/71 orally treated patients and 54/102 intravenously treated patients, with Influenza A being the most common pathogen in the oral group ($n = 3$), while Rhinovirus ($n = 4$) was most common pathogen in the intravenous group. Urine antigen tests were more often performed in intravenously treated patients (34% vs. 14%) (Appendices 12–19).

Antibiotic treatment

In the matched study population, amoxicillin (31/71), doxycycline (19/71), moxifloxacin (13/71) and amoxicillin clavulanic acid (6/71) were the most common antibiotics in the oral group; while amoxicillin (38/102), ceftriaxone (27/102), cefuroxime (15/102) and amoxicillin clavulanic acid (11/102) were most common in the intravenous group (Appendix 21). In the oral group, 5 out of 71 patients (7%) switched from oral to intravenous antibiotics (median switch time 1 day¹). In the oral group, 9 patients received a single intravenous dose at the ED (2 received amoxicillin, 6 ceftriaxone, and 1 ceftriaxone and tobramycin) before starting oral therapy. In the intravenous group, 84/102 patients (82%) switched to oral antibiotics (median switch time 2 days^{1,2,3}). Of these patients, 72/84 (86%) switched within three days (Appendix 23).

DISCUSSION

In this study, we explored the safety and efficacy of oral versus intravenous antibiotics as prescribed in current daily practice in patients with moderate-to-severe CAP (PSI class III–V) at the ED in a large teaching hospital in the Netherlands. We found no significant differences in 30-day mortality rate, ICU-admissions or readmissions. We observed that intravenous treatment was associated with an extra 2.6 days of admission (95% CI (1.2–4.0), $p < 0.001$) compared with oral treatment. This suggests that oral treatment, when based on clinical judgement of the ED-physician, might be a reasonable treatment option for patients with moderate-to-severe CAP.

The mortality rate for the complete (unmatched) cohort was higher than expected based on the PSI score, specifically for those treated intravenously. This is most likely the result of only including patients for whom a blood culture was performed. Conversely, orally treated patients showed a lower or similar mortality rate than expected. This difference may mainly reflect the clinical judgement of the treating

physician. For instance, the physician decided to treat the patient intravenously even when the PSI suggested moderate CAP. For the PSI class V patients who were treated orally, the assessment was made that oral treatment would be sufficient despite classification as severe CAP. This is also reflected in the differences in vital functions at baseline. Gut feeling, an important factor in clinical judgement, was found to be the second most important factor in identifying sepsis in a study that aimed to gain insight in the clinical decision making in a population of general practitioners.³⁰ This supports our hypothesis that clinical judgement is an important factor in assessing severity of illness and choosing the appropriate method of administration.

In the oral group, 31% were treated as outpatients. In times when healthcare systems have increasingly limited capacity for inpatient care, exploring treatment options that can be given outside the hospital are of great importance. Interestingly, we found no difference in LOS between oral and intravenous treatment for those admitted to the hospital. This might partly be explained by the fact that 82% of our intravenously treated patients were switched to oral treatment, of which 86% switched within three days, which is considered to be an early switch in literature.³¹ Such an early switch is associated with a reduction in length of hospital stay of two days in patients with severe CAP.³¹

In this study, 10 out of the 88 patients treated orally received one dose of intravenous antibiotics before starting oral treatment. As it is known that such a dose has a big impact on the bacterial load of sensitive micro-organisms, one might argue that our oral group is actually not a strictly oral group.³²⁻³⁴ We did allow for this approach, since one intravenous dose would still allow further treatment at home instead of hospital admission. The sub-analysis comparing only patients with strictly oral treatment versus intravenous treatment did not lead to different conclusions.

Oral treatment of CAP has already been widely evaluated and implemented for the treatment of mild-to-moderate cases.^{6-10,13} However, the role of oral antibiotics in the upfront treatment of more severe cases has not been sufficiently assessed despite the high bioavailability of oral antibiotics like amoxicillin, doxycycline and moxifloxacin.^{17-19,35} Only two studies with a large majority of moderate-to-severe CAP patients showed similar or even better outcomes for oral treatment compared to intravenous treatment, but did not specify the disease severity for patients with treatment failure.^{14,15} Furthermore, a large observational study in Canada showed that 13.7% of CAP PSI class IV and V patients could safely be managed as outpatients.¹⁶ Meanwhile, a recent large study showed that the vast majority of CAP patients are still treated with intravenous antibiotics, and early switch only

happened in 6% of the patients.³⁶ Our study adds to the evidence that the role of oral antibiotics could likely be increased in the management of moderate-to-severe-CAP.

A strength of this study is that it is an accurate representation of daily ED practice in a heterogeneous, real-life population. Furthermore, we believe that the quality of the data collection was excellent, with almost no missing data. However, our study also has limitations. First, this study describes a relatively small population limited to one hospital, which might lead to concerns regarding the power to detect relevant differences. Therefore, we emphasise that this is an exploratory study, and our results should not be seen as prove for non-inferiority of oral versus intravenous treatment. Furthermore, antibiotic treatment regimens will differ according to local epidemiology and resistance patterns, so extrapolating these results to other regions should be done with caution. Third, this study is susceptible to information bias, as we were only able to use the available information in the EHR. Hypothetically, although it would be unlikely to differ between the two groups, patients could have been readmitted to other hospitals without our knowledge. Although we applied state-of-the-art statistical methods such as propensity score matching and multiple imputation to balance differences between the orally and intravenously treated patients, there is a risk of residual confounding. We hypothesise that this residual confounding is mainly caused by clinical judgement, which would be acceptable since this study was designed to explore the outcomes of patient who were given oral treatment based on clinical judgement. However, there is a possibility that other factors have led to residual confounding that we have not taken into account. Finally, another limitation is the fact that we do not have a clear definition for clinical judgement, due to the retrospective nature of the study. We believe that clinical judgment or gut feeling plays a big role in identifying patients suitable for oral treatment, and future studies should carefully consider how to take clinical judgement into account.

CONCLUSION

Oral antibiotics appear to be a safe and effective treatment of moderate-to-severe CAP for selected patients based on clinical judgement of the attending physician. The dogma 'oral is the new iv'³⁷ should be the topic of future studies on moderate-to-severe CAP patients, with the aim to avoid overtreatment, and thereby improving antibiotic stewardship, and unnecessary hospitalisations.

Data availability

The datasets used in this study are not publicly available as they contain health related data. Limited datasets without any identifiable data are available from the corresponding author upon reasonable request.

Funding

The PredictED study was supported by unrestricted grant T18-040 from the Haga Scientific Research Fund. The funding source had no role in the design and conduct of the study.

Author information**Author notes**

These authors contributed equally: Anna G. Kaal and Rick Roos.

Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by A.K., R.R. and P.J. The first draft of the manuscript was written by A.K., R.R., and P.J. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethics declarations**Competing interests**

Author MN has an advisory role in the guidelines NAFLD/NASH and GORD and served as an unpaid QQ member for the Dutch Academic General Practice Network. All other authors have no relevant financial or non-financial interests to disclose.

SUPPLEMENTAL MATERIAL



REFERENCES

1. Wunderink, R. G. Community-acquired pneumonia. *Infectious Diseases*. 2017;251–257.e1. doi: 10.1016/B978-0-7020-6285-8.00028-9. Epub 2016 Aug 12. PMID: PMC7149822.
2. McLaughlin, J. M., Khan, F. L., Thoburn, E. A., Isturiz, R. E., Swerdlow, D. L. Rates of hospitalization for community-acquired pneumonia among US adults: A systematic review. *Vaccine*. 2020 Jan 22;**38(4)**:741-751. doi: 10.1016/j.vaccine.2019.10.101. Epub 2019 Dec 13. PMID: 31843272.
3. Welte, T., Torres, A., Nathwani, D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax*. 2012 Jan;**67(1)**:71-9. doi: 10.1136/thx.2009.129502. Epub 2010 Aug 20. PMID: 20729232.
4. Tsoumani, E., Carter, J. A., Salomonsson, S., Stephens, J. M., Bencina, G. Clinical, economic, and humanistic burden of community acquired pneumonia in Europe: a systematic literature review. *Expert Rev Vaccines*. 2023 Jan-Dec;**22(1)**:876-884. doi: 10.1080/14760584.2023.2261785. Epub 2023 Oct 13. PMID: 37823894.
5. Kolditz, M., Tesch, F., Mocke, L., Höffken, G., Ewig, S., Schmitt, J. Burden and risk factors of ambulatory or hospitalized CAP: A population based cohort study. *Respir Med*. 2016 Dec;**121**:32-38. doi: 10.1016/j.rmed.2016.10.015. Epub 2016 Oct 26. PMID: 27888989.
6. Metlay, J. P., Waterer, G. W., Long, A. C. et al. Diagnosis and treatment of adults with community-acquired pneumonia. an official clinical practice guideline of the american thoracic society and infectious diseases society of america. *Am J Respir Crit Care Med*. 2019 Oct 1;**200(7)**:e45-e67. doi: 10.1164/rccm.201908-1581ST. PMID: 31573350; PMID: PMC6812437.
7. Lim, W. S., Baudouin, S. V., George, R. C. et al. Pneumonia guidelines committee of the BTS standards of care committee. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*. 2009 Oct;**64 Suppl 3**:iii1-55. doi: 10.1136/thx.2009.121434. PMID: 19783532.
8. National Institute for Health and Care Excellence (NICE). *Pneumonia in Adults: diagnosis and management*. Published 3 December 2014, updated 7 July 2022.
9. National Institute for Health and Care Excellence (NICE). *Pneumonia (community-acquired): antimicrobial prescribing*. Published 16 September 2019.
10. Wiersinga, W. J., Bonten, M. J., Boersma, W. G. et al. Management of community-acquired pneumonia in adults: 2016 guideline update from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). *Neth J Med*. 2018 Jan;**76(1)**:4-13. PMID: 29380739.
11. Fine, M. J., Auble, T. E., Yealy, D. M. et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997 Jan 23;**336(4)**:243-50. doi: 10.1056/NEJM199701233360402. PMID: 8995086.
12. Lim, W. S., van der Eerden, M. M., Laing, R. et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003 May;**58(5)**:377-82. doi: 10.1136/thorax.58.5.377. PMID: 12728155; PMID: PMC1746657.

13. Teng, G. L., Chi, J. Y., Zhang, H. M., Li, X. P., Jin, F. Oral vs. parenteral antibiotic therapy in adult patients with community-acquired pneumonia: a systematic review and meta-analysis of randomized controlled trials. *J Glob Antimicrob Resist*. 2023 Mar;**32**:88-97. doi: 10.1016/j.jgar.2022.12.010. Epub 2023 Jan 18. PMID: 36669558.
14. Erard, V., Lamy, O., Bochud, P. Y., Bille, J., Cometta, A., Calandra, T. Full-course oral levofloxacin for treatment of hospitalized patients with community-acquired pneumonia. *Eur J Clin Microbiol Infect Dis*. 2004 Feb;**23**(2):82-8. doi: 10.1007/s10096-003-1060-x. Epub 2004 Jan 15. PMID: 14727149.
15. Marrie, T. J., Wu, L. Factors influencing in-hospital mortality in community-acquired pneumonia: a prospective study of patients not initially admitted to the ICU. *Chest*. 2005 Apr;**127**(4):1260-70. doi: 10.1378/chest.127.4.1260. PMID: 15821203; PMCID: PMC7095188.
16. Marrie, T. J., Huang, J. Q. Admission is not always necessary for patients with community-acquired pneumonia in risk classes IV and V diagnosed in the emergency room. *Can Respir J*. 2007 May-Jun;**14**(4):212-6. doi: 10.1155/2007/451417. PMID: 17551596; PMCID: PMC2676365.
17. Huttner A, Bielicki J, Clements MN, et al. Oral amoxicillin and amoxicillin-clavulanic acid: properties, indications and usage. *Clin Microbiol Infect*. 2020;**26**(7):871-879. doi:10.1016/j.cmi.2019.11.028. PMID: 31811919.
18. Öbrink-Hansen K, Hardlei TF, Brock B, et al. Moxifloxacin pharmacokinetic profile and efficacy evaluation in empiric treatment of community-acquired pneumonia. *Antimicrob Agents Chemother*. 2015;**59**(4):2398-2404. doi:10.1128/AAC.04659-14. PMID: 2566615;. PMCID: PMC4356773.
19. Van Den Broek AK, Visser CE, Veenstra J, Van Den Berg BTJ, Prins JM, Van Hest RM. The effect of the acute phase of infection on absorption of and exposure to orally administered antibiotics in non-critically ill, hospitalized patients. *J Antimicrob Chemother*. 2023;**78**(2):389-396. doi:10.1093/jac/dkac401. PMID: 36433818; PMCID: PMC9890209.
20. Hill SA, Jones KH, Lees LJ. Pharmacokinetics of parenterally administered amoxycillin. *J Infect*. 1980;**2**(4):320-332. doi:10.1016/s0163-4453(80)92720-6. PMID: 7185932.
21. Sutherland R, Croydon EA, Rolinson GN. Amoxycillin: a new semi-synthetic penicillin. *Br Med J*. 1972;**3**(5817):13-16. doi:10.1136/bmj.3.5817.13. PMID: 4402672; PMCID: PMC1788503.
22. de Velde F, Mouton JW, de Winter BCM, van Gelder T, Koch BCP. Clinical applications of population pharmacokinetic models of antibiotics: Challenges and perspectives. *Pharmacol Res*. 2018;**134**:280-288. doi:10.1016/j.phrs.2018.07.005. PMID: 30033398.
23. Charlson, M. E., Pompei, P., Ales, K. L., MacKenzie, C. R. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;**40**(5):373-83. doi: 10.1016/0021-9681(87)90171-8. PMID: 3558716.
24. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet*. 1974;**2**(7872):81-84. doi:10.1016/s0140-6736(74)91639-0. PMID: 4136544.
25. Azur, M. J., Stuart, E. A., Frangakis, C., Leaf, P. J. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res*. 2011 Mar;**20**(1):40-9. doi: 10.1002/mpr.329. PMID: 21499542; PMCID: PMC3074241.
26. Steyerberg, E. W. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer; 2009.

27. Zhao QY, Luo JC, Su Y, Zhang YJ, Tu GW, Luo Z. Propensity score matching with R: conventional methods and new features. *Ann Transl Med*. 2021;**9(9)**:812. doi:10.21037/atm-20-3998. PMID: 34268425; PMCID: PMC8246231.
28. Dregmans, E., Kaal, A. G., Meziyerh, S. et al. Analysis of variation between diagnosis at admission vs discharge and clinical outcomes among adults with possible bacteremia. *JAMA Netw Open*. 2022 Jun 1;**5(6)**:e2218172. doi: 10.1001/jamanetworkopen.2022.18172. PMID: 35737389; PMCID: PMC9226997.
29. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;**315(8)**:801-810. doi:10.1001/jama.2016.0287. PMID: 26903338; PMCID: PMC4968574.
30. Loots, F. J., Arpots, R., van den Berg, R., Hopstaken, R. M., Giesen, P., Smits, M. Recognition of sepsis in primary care: a survey among GPs. *BJGP Open*. 2017 May 3;**1(2)**:bjgpopen17X100965. doi: 10.3399/bjgpopen17X100965. PMID: 30564668; PMCID: PMC6169959.
31. Oosterheert, J. J., Bonten, M. J., Schneider, M. M. et al. Effectiveness of early switch from intravenous to oral antibiotics in severe community acquired pneumonia: multicentre randomised trial. *BMJ*. 2006 Dec 9;**333(7580)**:1193. doi: 10.1136/bmj.38993.560984.BE. Epub 2006 Nov 7. PMID: 17090560; PMCID: PMC1693658.
32. Harris AM, Bramley AM, Jain S, et al. Influence of Antibiotics on the Detection of Bacteria by Culture-Based and Culture-Independent Diagnostic Tests in Patients Hospitalized With Community-Acquired Pneumonia. *Open Forum Infect Dis*. 2017;**4(1)**:ofx014. Published 2017 Feb 10. doi:10.1093/ofid/ofx014. PMID: 28480285; PMCID: PMC5414111.
33. Pertel PE, Bernardo P, Fogarty C, et al. Effects of prior effective therapy on the efficacy of daptomycin and ceftriaxone for the treatment of community-acquired pneumonia. *Clin Infect Dis*. 2008;**46(8)**:1142-1151. doi:10.1086/533441. PMID: 18444848.
34. Driscoll AJ, Deloria Knoll M, Hammitt LL, et al. The Effect of Antibiotic Exposure and Specimen Volume on the Detection of Bacterial Pathogens in Children With Pneumonia. *Clin Infect Dis*. 2017;**64(suppl_3)**:S368-S377. doi:10.1093/cid/cix101. PMID: 28575366; PMCID: PMC5447850.
35. Agwuh KN, MacGowan A. Pharmacokinetics and pharmacodynamics of the tetracyclines including glycylcyclines. *J Antimicrob Chemother*. 2006;**58(2)**:256-265. doi:10.1093/jac/dkl224. PMID: 16816396.
36. Deshpande, A., Klompas, M., Guo, N. et al. Intravenous to oral antibiotic switch therapy among patients hospitalized with community-acquired pneumonia. *Clin Infect Dis*. 2023 Jul 26;**77(2)**:174-185. doi: 10.1093/cid/ciad196. PMID: 37011018; PMCID: PMC10527888.
37. Davar, K., Clark, D., Centor, R. M. et al. Can the future of ID escape the inertial dogma of its past? The exemplars of shorter Is better and oral Is the new IV. *Open Forum Infect Dis*. 2022 Dec 29;**10(1)**:ofac706. doi: 10.1093/ofid/ofac706. PMID: 36694838; PMCID: PMC9853939.



Chapter 9

Choose oral over intravenous antibiotic therapy

Anna G. Kaal, Rick Roos, Erik B. Wilms, Hanneke Borgdorff, Cees van Nieuwkoop

ABSTRACT

Choose oral over intravenous antibiotic therapy

Many physicians and patients hold the (unconscious) belief that intravenous antibiotic therapy is superior to oral therapy. This belief is also reflected in guidelines, where increasing severity of infection often leads to the recommendation of intravenous options only. But is this belief justified, and where does it come from? Treating with oral antibiotics has many potential advantages, such as fewer hospital admissions and the prevention of related complications. In this era of increasing demand for appropriate and efficient care, oral antibiotic treatment should replace intravenous treatment where possible. Here, we outline the crucial factors that should be considered when deciding between oral and intravenous treatment of infections: pharmacokinetics and dynamics, antibiotic resistance, and specific patient factors.

Veel artsen denken – bewust of onbewust – dat je antibiotica beter intraveneus dan oraal kunt geven. Ook richtlijnen adviseren bij ernstige ziekte vaak alleen nog intraveneuze behandeling. Onterecht, vinden wij. Behandel bij voorkeur oraal.

Antibiotische behandeling via de orale weg heeft potentieel veel voordelen, zoals minder ziekenhuisopnames en minder complicaties als gevolg daarvan. In deze tijd waarin de roep om passende en doelmatige zorg steeds luider wordt, moet orale antibiotische behandeling de intraveneuze therapie vervangen waar dit kan.

In dit artikel schetsen wij de cruciale factoren die aan bod dienen te komen bij de keuze tussen orale en intraveneuze behandeling van infecties: farmacokinetiek en -dynamiek, resistentie en specifieke patiëntfactoren en factoren in de zorgketen.

Hoe werd intraveneuze therapie de norm?

Toen in de jaren 30 en 40 van de vorige eeuw de eerste soorten antibiotica beschikbaar kwamen in de vorm van prontosil rubrum en penicilline, betekende dit een grote doorbraak in de behandeling van infectieziekten.¹ In die tijd waren er nog geen gerandomiseerde klinische trials (RCT's) en veel behandelregimes waren gebaseerd op empirische behandelingen en 'expert opinion'. Deze veelal zeer effectieve intraveneuze regimes werden in de loop der jaren de norm.¹ Hierdoor heeft de overtuiging dat intraveneuze therapie beter is dan orale therapie bij de behandeling van infectieziekten zich diep kunnen wortelen.

In de afgelopen decennia zijn steeds meer onderzoeken verschenen die aanleiding geven om deze voorkeur te heroverwegen. Het is echter niet makkelijk om de huidige antibioticarijchlijnen aan te passen. Paradoxaal genoeg is er hooggradig bewijs nodig om richtlijnen te veranderen, terwijl er lang niet altijd bewijs is voor de huidige aanbevelingen. Soms ontbreekt zelfs de pathofysiologische of farmacologische ratio erachter.

In een recente meta-analyse naar de antibiotische behandeling van patiënten met endocarditis, osteomyelitis of bacteriëmie bleek dat orale therapie non-inferieur is aan intraveneuze therapie, met minder therapiegerelateerde complicaties en een kortere opnameduur.² Voor de behandeling van urineweginfecties en community-acquired pneumonieën zijn vergelijkbare conclusies getrokken.^{3,4} Daarom wordt momenteel in verschillende richtlijnen van de Stichting Werkgroep Antibioticabeleid (SWAB) geadviseerd om bij een stabiele klinische situatie en een goede orale behandeloptie de intraveneuze behandeling na 48-72 uur om te zetten naar oraal.⁵ Het is echter belangrijk om te erkennen dat er geen rationale is waarom dit pas na 48 uur zou kunnen, en ook niet waarom dit uitsluitend zou kunnen bij een stabiele klinische situatie.⁶ Wij stellen dat de enige juiste voorwaarde is dat ter plaatse van de infectie een adequate concentratie van een geschikt antibioticum kan worden bereikt.

Direct beginnen met orale therapie

Er is relatief weinig literatuur over direct beginnen met orale behandeling van matig-ernstige infectieziekten, ook al wordt dit in de praktijk regelmatig toegepast, met name in de eerstelijnszorg. In een recent onderzoek onder patiënten met matig-ernstige pneumonie op de Spoedeisende Hulp vonden wij geen verschil in uitkomsten tussen orale en intraveneuze behandeling; alleen de opnameduur was langer bij patiënten die intraveneus waren behandeld.⁷ Een Nederlandse cohortstudie waarin patiënten met een urineweginfectie met weefselinvasie, waaronder urosepsis, op

basis van een pluis- of niet-pluisgevoel thuis oraal werden behandeld, liet ook geen verschil zien in klinische uitkomsten.⁸

Gezien het gebrek aan literatuur is het niet verbazingwekkend dat direct oraal behandelen zelden een optie is in de huidige tweedelijsrichtlijnen en dat artsen daarom terughoudend zijn. Maar doen wij onze patiënten daarmee niet te kort? Met de potentiële voordelen van orale behandeling in het achterhoofd, zoals meer mogelijkheden voor thuisbehandeling, pleiten wij voor meer aandacht voor orale behandeling van infectieziekten. Denk hierbij aan orale behandeling van ernstige pneumonie, urosepsis en cellulitis.

Naast de beschikbare literatuur zijn er nog drie andere aspecten belangrijk om te overwegen bij de keuze tussen orale of intraveneuze behandeling van een infectieziekte: (a) farmacokinetiek en -dynamiek; (b) resistentie; en (c) relevante factoren betreffende de patiënt en de zorgketen.

Farmacokinetiek en -dynamiek

Een belangrijk verschil tussen orale en intraveneuze toediening van geneesmiddelen is de farmacokinetiek. De biologische beschikbaarheid van intraveneuze middelen is 100%, terwijl die van orale middelen afhankelijk is van de absorptie. Goede absorptie vereist een zekere mate van oplosbaarheid in water (oplossen in darminhoud) en oplosbaarheid in vet (snelle diffusie). Antibiotica die alleen wateroplosbaar zijn, zoals aminoglycosiden, benzylpenicilline en sommige cefalosporines – onder andere cefuroxim en ceftriaxon –, worden oraal onvoldoende geabsorbeerd. Om die reden worden deze middelen veelal intraveneus toegediend.

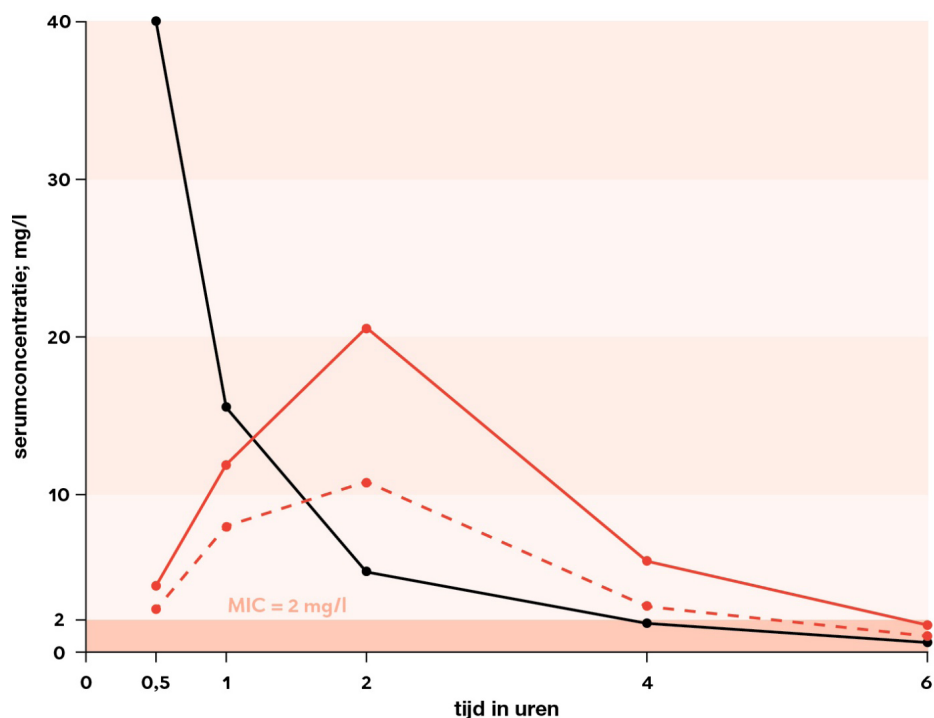
In de tabel staat een overzicht van antibiotica met een goede biologische beschikbaarheid bij orale toediening; het moet worden overwogen om deze middelen direct oraal toe te dienen, mits er geen aanwijzingen zijn voor malabsorptie.⁹ Bij het voorschrijven van orale middelen moet rekening worden gehouden met andere geneesmiddelen en voeding die de absorptie kunnen beïnvloeden, zoals zuivelproducten bij ciprofloxacin en eventueel 'over the counter' verkregen antacida bij doxycycline.

antibiotica met goede biologische beschikbaarheid na orale inname**amoxicilline*****ciprofloxacin****clindamycine****cotrimoxazol****doxycycline****metronidazol****moxifloxacin**

* Orale toediening van amoxicilline wordt enkel geadviseerd bij behandeling van micro-organismen met een lage minimaal inhiberende concentratie, zoals streptokokkenspecies, en bij infecties buiten het centrale zenuwstelsel.

Een voorbeeld: amoxicilline

De rol van farmacokinetiek en -dynamiek leggen we uit aan de hand van het veelgebruikte antibioticum amoxicilline. Bij intraveneuze toediening wordt direct na toediening een hoge piekconcentratie bereikt, gevolgd door een snelle daling (korte halfwaardetijd). Bij orale toediening gaat de opname geleidelijk en wordt de piekconcentratie later bereikt, en gaat de daling minder snel door voortdurende resorptie ('flip-flop'-fenomeen; figuur).^{10,11} Dit zorgt ervoor dat de concentratie gedurende een langere periode hoger is dan bij intraveneuze toediening van amoxicilline. De effectiviteit van dit antibioticum wordt namelijk met name bepaald door de tijd dat de concentratie in het bloed hoger is dan de minimaal inhiberende concentratie (MIC) van het micro-organisme. Bij orale toediening is de concentratie van amoxicilline zelfs gedurende een langere tijd boven de MIC dan bij intraveneuze toediening.



MIC = minimaal inhiberende concentratie.

amoxicilline: — 1000 mg intraveneus

— 1000 mg oraal

- - - 500 mg oraal

MIC 95%-betrouwbaarheidsinterval	
Streptococcus pneumoniae	0,016-0,06 mg/l
Haemophilus influenzae	0,5-2,0 mg/l

Er kan dus worden geconcludeerd dat orale toediening van amoxicilline geschikt is, mits de absorptie adequaat is. De locatie van de infectie speelt hierbij wel een belangrijke rol. Bij een meningitis bijvoorbeeld is een hogere concentratie vereist om de bloed-hersenbarrière te passeren. Zo'n hoge concentratie kan alleen met intraveneuze toediening worden bereikt.

Resistentie

Bij de keuze voor een antibioticum moet tevens rekening worden gehouden met mogelijke resistentie van de meest waarschijnlijke ziekteverwekkers. Ciprofloxacin is bijvoorbeeld de eerste keuze bij de orale behandeling van een patiënt met een

urineweginfectie en koorts. Daarbij is de meest waarschijnlijke verwekker, *E. coli*, in 16% van de gevallen resistent; voor ceftriaxon is dit percentage 7%.¹² Het klinisch beeld bepaalt welk risico op resistentie acceptabel is.

Patiënt en de zorgketen

Doordat intraveneuze therapie al snel gelijkstaat aan een ziekenhuisopname, is het belangrijk af te wegen of de verwachte voordelen opwegen tegen eventuele schade aan de patiënt. Denk hierbij aan kwetsbare ouderen bij wie het voorkómen van ziekenhuiscomplicaties, zoals een delier, essentieel is. Daarnaast kunnen psychosociale factoren spelen die een ziekenhuisopname onwenselijk maken.

Een cruciale factor bij orale behandeling van matig-ernstige infecties in de thuissetting is dat de veiligheid gewaarborgd moet zijn. Dit omvat de aanwezigheid van thuiszorg of mantelzorg, afspraken over follow-up en bereikbaarheid van (specialistische) zorg indien nodig, en een tijdige overdracht aan de huisarts. Zeker nu de eerstelijnszorg onder druk staat vanwege personeelstekort en een hoge werkdruk, is het niet vanzelfsprekend dat de huisarts patiënten die thuis worden behandeld, intensief kan controleren. Om de veiligheid te waarborgen kunnen bijvoorbeeld ketenzorgpaden worden ontwikkeld en geëvalueerd, zoals we nu doen in de regio Haaglanden voor de behandeling van matig-ernstige luchtweginfecties.¹³

CONCLUSIE

Op dit moment is er op basis van richtlijnen weinig ruimte voor orale behandeling van matig-ernstige infectieziekten, en onderzoek naar orale behandelopties is hard nodig. Door ons bewust te zijn van het gebrek aan bewijs voor superioriteit van intraveneuze behandeling en door farmacokinetische en -dynamische aspecten van orale behandelopties en relevante patiëntfactoren mee te wegen, zien we waarschijnlijk meer ruimte voor orale behandeling dan tot nu toe gedacht.

Belangenverstrengeling

Belangenconflict en financiële ondersteuning: geen gemeld.

Verantwoording

Anna Kaal en Rick Roos hebben een gelijke bijdrage geleverd aan dit artikel en zijn daarom gedeeld eerste auteur.

LITERATUUR

1. Davar K, Clark D, Centor RM, et al. Can the future of ID escape the inertial dogma of its past? The exemplars of shorter is better and oral is the new IV. *Open Forum Infect Dis.* 2022;10(1):ofac706. doi:10.1093/ofid/ofac706. Medline
2. Wald-Dickler N, Holtom PD, Phillips MC, et al. Oral Is the New IV. Challenging decades of blood and bone infection dogma: a systematic review. *Am J Med.* 2022;135(3):369-379. e1. doi:10.1016/j.amjmed.2021.10.007. Medline
3. Athanassa Z, Makris G, Dimopoulos G, Falagas ME. Early switch to oral treatment in patients with moderate to severe community-acquired pneumonia: a meta-analysis. *Drugs.* 2008;68(17):2469-2481. doi:10.2165/0003495-200868170-00005. Medline
4. Pohl A. Modes of administration of antibiotics for symptomatic severe urinary tract infections. *Cochrane Database Syst Rev.* 2007;2007(4):CD003237. doi:10.1002/14651858.CD003237.pub2. Medline
5. SWAB Guidelines for Antimicrobial Stewardship. <https://swab.nl/en/swab-guidelines>, geraadpleegd op 22 mei 2024.
6. Van Den Broek AK, Visser CE, Veenstra J, Van Den Berg BTJ, Prins JM, Van Hest RM. The effect of the acute phase of infection on absorption of and exposure to orally administered antibiotics in non-critically ill, hospitalized patients. *J Antimicrob Chemother.* 2023;78(2):389-396. doi:10.1093/jac/dkac401. Medline
7. Kaal AG, Roos R, de Jong P, et al. Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study. *Sci Rep.* 2024;14(1):8271. doi:10.1038/s41598-024-59026-2. Medline
8. Van Nieuwkoop C, van 't Wout JW, Spelt IC, et al. Prospective cohort study of acute pyelonephritis in adults: safety of triage towards home based oral antimicrobial treatment. *J Infect.* 2010;60(2):114-121. doi:10.1016/j.jinf.2009.11.008. Medline
9. Sevinç F, Prins JM, Koopmans RP, Langendijk PN, Dankert J, Speelman P. Vroege omzetting van intraveneuze naar orale antibiotica: 'switchtherapie'. *Ned Tijdschr Geneesk.* 1999;143(47):2364-2369. Medline.
10. Sutherland R, Croydon EA, Rolinson GN. Amoxycillin: a new semi-synthetic penicillin. *BMJ.* 1972;3(5817):13-16. doi:10.1136/bmj.3.5817.13. Medline
11. Hill SA, Jones KH, Lees LJ. Pharmacokinetics of parenterally administered amoxycillin. *J Infect.* 1980;2(4):320-332. doi:10.1016/S0163-4453(80)92720-6. Medline
12. De Greef SD, Kolwijck E, Schoffelen AF. NethMap 2023. Consumption of antimicrobial agents and antimicrobial resistance among medically important bacteria in the Netherlands in 2022 / MARAN 2023. Monitoring of Antimicrobial Resistance and Antibiotic Usage in Animals in the Netherlands in 2022. Bilthoven: RIMV; 2023.
13. Roos R, Pepping RMC, van Aken MO, et al. Evaluation of an integrated care pathway for out-of-hospital treatment of older adults with an acute moderate-to-severe lower respiratory tract infection or pneumonia: protocol of a mixed methods study. *BMJ Open.* 2023;13(8):e073126. doi:10.1136/bmjopen-2023-073126. Medline

Choose oral over intravenous antibiotic therapy



Chapter 10

General discussion

GENERAL DISCUSSION

The overall aim of this thesis was to identify and evaluate strategies to improve the management of infectious diseases in adult ED patients. Diagnostic testing and treatment should be provided to patients who require them, while minimizing unnecessary interventions. This chapter first states the key findings of this thesis, followed by a more detailed discussion of all chapters, including future directions and broader considerations. This is done according to the three main parts: diagnostics, risk stratification and treatment.

KEY FINDINGS

- We developed multiple prediction models to improve the diagnostic process of bacteremia, aiming to safely reduce the number of blood cultures in patients suspected of community-acquired bacteremia at the ED. Our most promising model was able to save around 30% of blood cultures, while missing 1-5% of true bacteremia cases. We showed that especially procalcitonin was strongly predictive for true bacteremia across multiple large international cohorts.
- The second part of this thesis focused on risk prediction. First, we showed that misdiagnosis occurs in 1 in 9 patients suspected of infection at the ED and is associated with more exposure to broad-spectrum antibiotics but not with other severe outcomes. We identified risk factors for misdiagnosis, of which a positive urinary dip stick without any symptoms relatable to a urinary tract infection was most remarkable. Second, we compared the two most commonly used prediction models to guide treatment decisions for community-acquired pneumonia and showed that the use of the CURB-65 was associated with reduced mortality compared to the PSI score.
- Finally, in the third part of this thesis we compared oral versus intravenous antibiotics in patients with moderate-to-severe community-acquired pneumonia (CAP) and found no differences in outcomes for both treatment strategies. This finding forms a much needed basis for conducting a randomized controlled trial on oral versus intravenous treatment of CAP.

PART I DIAGNOSTICS

In our search for improvement of diagnostics in infectious diseases we focused on the diagnostic process of bacteremia. Bacteremia, defined as the presence of bacteria in the bloodstream, is associated with substantial morbidity and mortality.¹ The current diagnostic process will typically begin in the emergency department (ED), where patients present with signs of infection, ranging from non-specific

symptoms such as fever, to severe hypotension indicative of sepsis. Clinicians assess the risk of bacteremia and, if warranted, obtain blood cultures.²

These blood cultures consist of two bottles that are filled with patient blood via venipuncture. These cultures are incubated in the microbiology laboratory and continuously monitored for bacterial growth. When bacteremia is present, the first signals of positivity generally appear within 24 hours.³ Subsequent pathogen identification and susceptibility testing require additional time, resulting in a diagnostic delay of several days before final results become available.

There are multiple aspects of this process that need improvement: To minimize the risk of missing bacteremia, clinicians often obtain blood cultures liberally, yielding a low diagnostic return: only 4–11% of cultures are positive.^{4–6} Meanwhile, blood cultures are costly: ranging from 96 to 423 USD per set.⁷ Moreover, when a blood culture is not collected properly, skin commensals can be introduced into the culture bottles, falsely leading to a positive blood culture result: contamination. Of all blood cultures taken at the ED, 2–8% are contaminated.^{4–9} It can be hard to determine whether a positive blood culture results reflects true bacteremia or contamination. Consequently, contaminated blood cultures are associated with increased length of stay, more exposure to antibiotics, higher costs, and even increased mortality.^{7–9} It would therefore be of great value to find a way to identify patients with a low risk of bacteremia, in whom we can safely withhold blood cultures.

In **Chapter 2** we showed that procalcitonin (PCT) can be useful to exclude community-acquired bacteremia (CAB).¹⁰ PCT is a biomarker that has received much interest because, unlike commonly used inflammatory markers such as C-reactive protein (CRP), it more specifically reflects bacterial rather than viral infections. This is because PCT levels rise in response to pro-inflammatory cytokines and bacterial toxins, while it is suppressed by higher levels of interferon- γ due to viral infections.¹¹

We performed a systematic review and meta-analysis and found that PCT at a low threshold (0.10 ng/mL) has high sensitivity for CAB. This implies that low values can be used to rule out bacteremia, and therefore, can support withholding blood cultures. Notably, especially the studies with low risk of bias showed high sensitivity (97%, 95% CI: 92–99%). However, we observed considerable heterogeneity in study methodology: 15 of the 40 included studies either lacked a clear definition of blood culture contamination or excluded such cultures from analysis altogether. Studies applying a clear definition demonstrated higher sensitivity, which is biologically plausible since contaminated cultures do not reflect true infection and therefore typically do not induce elevated PCT levels. If these cultures are analyzed as true positives, this would falsely lead to a reduced sensitivity. This finding underlines the need for consensus on an internationally accepted definition of contaminated blood

cultures. A Delphi process could be used to establish such consensus, building on existing criteria such as those proposed by the College of American Pathologists (CAP), the Clinical and Laboratory Standards Institute (CLSI), or the approach we developed in Chapter 4.^{12,13}

After two previous meta-analysis regarding diagnostic accuracy, this meta-analysis added relevant knowledge about diagnostic properties of PCT at lower thresholds.^{14,15} While some might find the associated sensitivity acceptable for use in clinical practice, others might not, as it is known that a considerable number of clinicians require a sensitivity for bacteremia of 99-100%.¹⁶ We conclude that additional studies on the diagnostic accuracy of PCT for CAB are unlikely to yield further insights. Future research should instead focus on combining PCT with clinical parameters to evaluate whether this would further improve diagnostic accuracy, and hence, clinical applicability.

This question was the starting point for **Chapter 3**. We developed multiple prediction models using combinations of PCT, the more commonly used C-reactive protein (CRP) and clinical characteristics to identify patients at low risk for CAB.¹⁷ In the development of these models, we especially focused on parameters that would be easy to implement in electronic health records. The models were derived off a Dutch dataset of the Haga Teaching Hospital including 2111 patients, and validated in a Spanish dataset of 4436 patients. The most promising model was able to save 29% of all blood cultures while only missing 1% of bacteremic episodes. However, calibration (is the estimated risk the same as the observed risk?), was suboptimal for all models that contained PCT. Therefore, we updated the models based on both the Dutch and Spanish dataset and performed external validation in a new setting.¹⁸

Chapter 4 describes this validation in a large US hospital in Ann Arbor, Michigan. Using retrospectively collected data from more than 12.000 patients, we showed that the updated models had the potential to reduce blood culture testing by 30% of blood cultures, missing 5% of true bacteremia episodes, with substantially improved calibration (**Figure 1**).

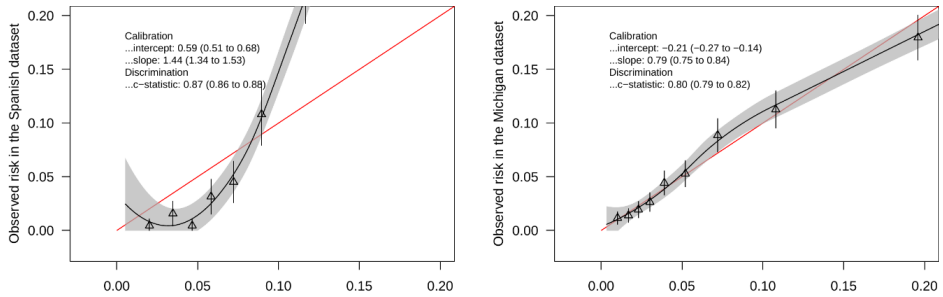


Figure 1. Calibration plots for the basic model with procalcitonin (PCT) in the Spanish data, and for the updated model in Michigan. The X-axis represents the predicted risks, the Y-axis shows the observed risks. Ideally, the calibration curve follows the red line indicating that predicted risks are in line with observed risks. We observe substantially improved calibration of the updated model.

There were multiple challenges in the development and validation of these models. First, we observed geographical variation in model performance, especially due to variation in the relation between procalcitonin and bacteremia risk. Multiple factors may have contributed to this variability.¹⁹ When performing an external validation, it can be challenging to acquire high quality data. Local definitions might differ from the original dataset, and retrospectively collected data might have limited information. For instance, deciding which blood cultures in the Michigan dataset were contaminated was challenging as this was not automatically retrievable. We worked together with clinicians from multiple disciplines to create a step-wise approach to classify blood cultures, but as we specifically wanted to prevent misclassification of true bacteremia as contamination we erred on the side of caution in case of doubt. This might have led to an overestimation of the number of true positive blood cultures. Furthermore, differences between healthcare systems can result in differences in case-mix, and differences in patient selection by using different criteria for PCT or blood culture testing might result in different risks of bacteremia.¹⁹ We observed a mean risk for bacteremia of 20% in the Spanish data, opposed to 9% in the Michigan data. However, we cannot exclude that there is a true difference in the relation between PCT in CAB in the Spanish population compared to the Dutch and American population. Further validation in more settings could increase our understanding of this geographical variation, although we believe this to be not essential before moving forward towards the next step.

When developing a clinical prediction model typically involved four steps: Development, validation, impact assessment and implementation (**Figure 2**).²⁰ Yet, most prediction models never make it to the implementation phase.²¹ We addressed the first two steps in chapters 3 and 4, but an impact study is essential for successful implementation.

An impact study that prospectively evaluates clinical outcomes when using the model, could answer several relevant questions. First: how should such a model be integrated in clinical practice? Should we provide hard-stops of blood culture orders based on a low risk prediction or should it be possible for the clinician to overrule the model? If overruling the model would be possible, how would that impact the model's potential to reduce blood cultures? Second, what would happen if the models are applied to patients who currently would not be tested for bacteremia, and how can such unintended use be prevented? Finally, we expect that the risk prediction will be available for the majority of patients at the moment that antibiotic therapy is started. However, some will receive antibiotic therapy sooner, warranting a different approach. In those patients, cultures could be drawn and the decision to move forward with analysis could be made based on the final risk prediction.

Additionally, a survey among clinicians could facilitate implementation by identifying barriers and clarifying the requirements for adopting a bacteremia risk model in practice. It is especially important to understand clinicians' perspectives on the acceptable risk threshold for withholding blood cultures. While the model provides an estimated bacteremia risk for each individual patient, the decision regarding what risk is considered sufficiently low to withhold blood cultures ultimately lies with the user. We used a 5% risk threshold to report our findings, but some might only accept a 1% risk. This would miss less true bacteremia episodes, but would automatically also reduce the number of saved blood cultures.

There will be multiple challenges regarding the implementation of our bacteremia risk models. Many clinicians may be unaware of the high costs and negative consequences of contaminated blood cultures, which could limit willingness to change current practice. In addition, PCT is a biomarker that is not widely available, and costs of procalcitonin and blood cultures vary substantially around the world.^{7,22,23} This complicates evaluation of cost-effectiveness. As a result, implementation may initially only be feasible in hospitals that already use PCT for patients with suspected CAB. Furthermore, in some hospitals, PCT results require multiple hours to become available, also limiting applicability. In such settings, other approaches to reduce blood cultures such as indication-based algorithms might be more feasible.²⁴

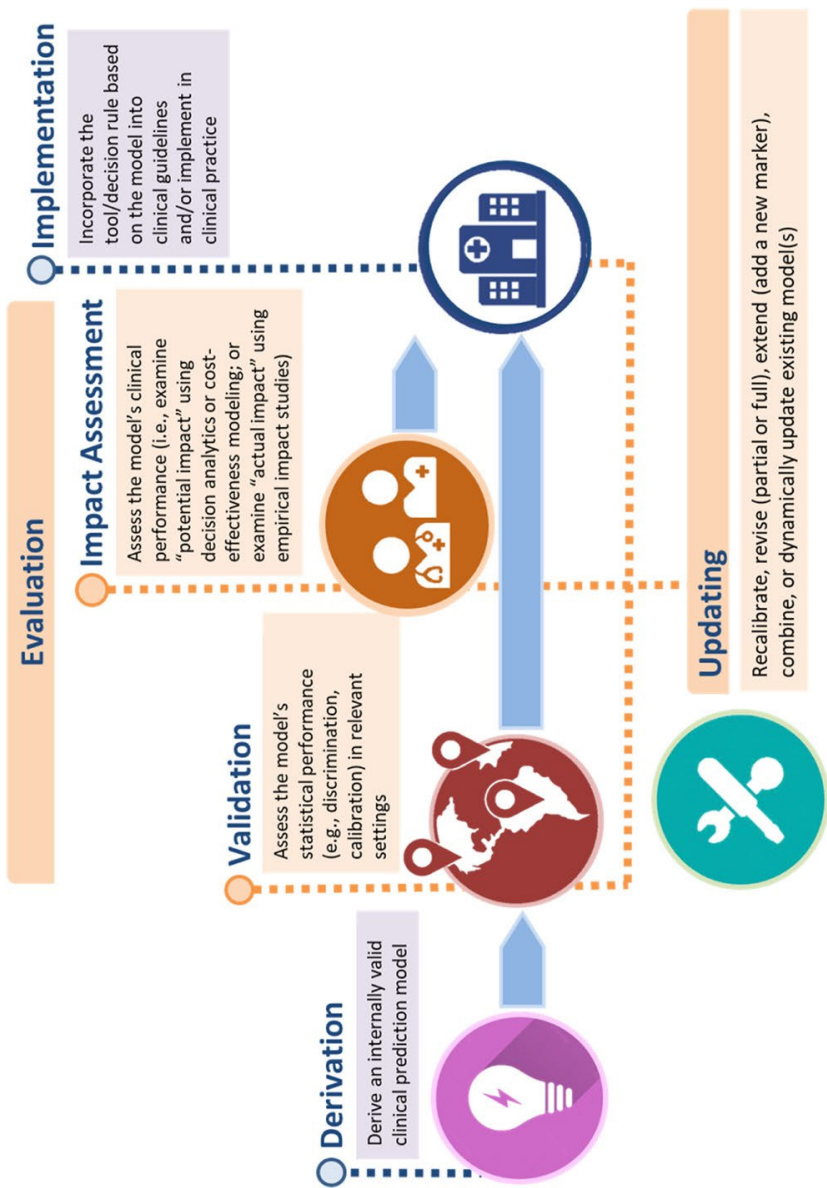


Figure 2: Steps in the development and implementation of clinical prediction models. Figure adapted from Binuya et al, 2022.²⁰

In summary, we have shown that depending on the acceptable risk threshold for bacteremia, PCT alone can help to reduce blood cultures use at the ED. Moreover, our models that combine PCT with clinical characteristics offer further improved diagnostic accuracy and are promising to optimize the diagnostic process of bacteremia.

Part II Risk stratification

Pressure on healthcare systems is increasing due to rising costs, shortages in healthcare staff, and, in some settings, limited resources.^{25,26} Identifying patients at high risk for severe outcomes can help prioritize care for those who need it most. Conversely, recognizing patients at low risk can prevent unnecessary exposure to treatments with potential harms. For example, while ICU admission can be life-saving, it also carries risks such as hospital-acquired infections or delirium.

This part of the thesis focused on risk stratification for poor outcomes in patients presenting to the ED with suspected infection. We approached this by identifying risk factors for adverse outcomes and by developing prognostic prediction models for a broad population of patients with suspected infection.

Because misdiagnosis in general patient populations is associated with worse outcomes, we hypothesized this might also apply to infectious diseases, particularly since treatment decisions are often guided by the most likely causal pathogens, which differ depending on the suspected site of infection.²⁷⁻²⁹ Before our study, no data were available on misdiagnosis rates in this setting. In **Chapter 5**, we showed that one in nine ED patients with suspected infection was misdiagnosed.³⁰ Surprisingly, we found no association with severe outcomes such as longer length of stay, ICU admission, or mortality. A possible explanation is that these patients were more frequently treated with broad-spectrum antibiotics, possibly reflecting diagnostic uncertainty among clinicians. Nevertheless, there may be negative consequences of misdiagnosis that we could not capture in this study. A prognostic model did identify patient characteristics associated with misdiagnosis, but as the performance was only fair it will likely not directly benefit clinicians. Risk factors included older age, dementia and a positive urine dipstick without signs of a urogenital tract infection. Especially this last risk factor could be an interesting starting point for future research. Reducing uncertainty regarding urinary tract infection diagnosis might prevent unnecessary exposure to broad-spectrum antibiotics.

For community-acquired pneumonia, prognostic models are already widely used. The Dutch national guideline recommends either the PSI score or the CURB-65 for risk stratification to guide treatment and site-of-care decisions.³¹⁻³³ Both models estimate 30-day mortality risk and classify patients into different risk groups. Most

studies suggest similar prognostic performance, although some report that the PSI score is better at identifying low-risk patients.³⁴⁻³⁷ This might explain why some clinicians prefer the PSI, while others favor the simpler and more user-friendly CURB-65.

In **Chapter 6**, we collaborated with the National Healthcare Institute to create a nationwide cohort of CAP patients using claims data.³⁸ We compared outcomes between hospitals using CURB-65 and those using the PSI score. Surprisingly, we found that in 50,984 patients, there was a lower mortality in 'CURB-65' hospitals. This difference might be explained because the CURB-65 is known to classify a larger proportion of patients as having 'severe pneumonia'.^{37,39,40} According to the national guideline, these patients should receive broad-spectrum antibiotics such as second- or third-generation cephalosporins, whereas less severe categories are treated with narrower-spectrum regimens.³¹ Therefore, some patients not classified as 'severe' by the PSI score might have been undertreated. However, given the retrospective design and limitations in claims data, residual confounding cannot be ruled out. Therefore, these findings are not sufficient to change guidelines. However, they should warrant an RCT comparing clinical outcomes according to use of CURB-65 versus the PSI score.

Finally, **Chapter 7** describes clinical and microbiological outcomes for patients with COVID-19 at the ED.⁴¹ When the COVID-19 pandemic started, hospitals were overwhelmed by the number of patients. We aimed to identify risk factors for severe disease to help allocate scarce resources. Furthermore, since little was known about the disease course, we evaluated bacterial coinfections. We found these to be rare (1.4%) and concluded that blood cultures have limited value in this population. Both procalcitonin and the PSI score, however, proved useful for risk stratification in patients with COVID-19. While these findings are less relevant today due to changes in the patient population, driven by factors such as immunization and genetic evolution of the virus itself, they illustrate the importance of rapid data collection.⁴² Because of the ongoing PredictED study during the pandemic, we were able to generate and share clinically relevant findings quickly, enabling a more informed response to a rapidly changing situation.

PART III TREATMENT

After diagnostics and risk stratification, treatment forms the next crucial step in infection management. In addition to selecting the appropriate medication to start, it is also important to decide on the route of administration. Most antibiotics will be administered either orally or intravenously (IV). In this part of the thesis, we explored the potential for oral antibiotic therapy in situations where IV antibiotics

are traditionally used. Oral treatment could reduce hospital admissions and prevent complications associated with IV administration, such as phlebitis and, more rarely, secondary infections like bacteremia or even endocarditis.

Community-acquired pneumonia (CAP) is one of the most common infectious diseases at the ED.⁴³⁻⁴⁵⁻³⁹ The Dutch guideline advises treatment with IV amoxicillin in moderate-to-severe pneumonia as classified by the PSI score or the CURB-65.³¹⁻³³ However, based on pharmacokinetic profiles, the local guideline of the Haga Teaching hospital differed from the national guideline with regard to this specific patient category. This local guideline advised oral amoxicillin if deemed appropriate by the treating clinician. This difference created a unique opportunity to compare clinical outcomes between oral and IV treatment in moderate-to-severe CAP, as described in **Chapter 7**.⁴⁶

We found no differences in clinical outcomes between the groups, except for a 2.5-day longer length of stay in the IV group, explained by higher admission rates. This study is among the first to evaluate immediate oral therapy in moderate-to-severe CAP. Given the high prevalence of CAP, changing first-line therapy from IV to oral antibiotics could significantly impact healthcare utilization and costs. Our findings provide a much-needed basis to continue with a randomized controlled trial studying clinical outcomes for oral versus IV treatment in moderate-to-severe CAP. If an RCT would confirm our findings, we believe this would be sufficient to change existing guidelines.

Interestingly, while current guidelines favor IV antibiotics, the supporting evidence is limited.^{31,47,48} Meanwhile, high levels of evidence are required in order to change guidelines. In the opinion article in **Chapter 8** we discussed why clinicians have come to believe that intravenous therapy is superior to oral therapy.⁴⁹ Furthermore, we explain why we believe that in deciding on route of administration, the only question should be whether an appropriate oral antibiotic can achieve adequate therapeutic concentrations at the site of infection. We describe different relevant aspects when considering oral treatment of infectious diseases. The paper outlines relevant considerations for oral antibiotic use and challenges existing assumptions, hopefully encouraging further research into oral treatment strategies for infectious diseases.

Main findings and recommendations	
Part 1: Diagnostics	
• Our bacteremia prediction model was able to save around 30% of blood cultures, while missing 1-5% of true bacteremia cases. Especially procalcitonin was highly predictive of bacteremia.	• A survey study could increase our understanding of how such models should be implemented. Impact studies are needed to facilitate successful implementation in clinical practice. • Further validation may serve as continuous monitoring of performance with potential for dynamic and local updating
Part 2: Risk stratification	
• 1 in 9 patients suspected of infection at the ED was misdiagnosed, without associated severe outcomes except for increased exposure to broad-spectrum antibiotics. Multiple risk factors for misdiagnosis were identified, such as a positive urinary dip stick without urinary symptoms	• Studies aimed at reducing uncertainty regarding urinary tract infection diagnosis might prevent unnecessary exposure to broad-spectrum antibiotics.
• In a nationwide study, the use of the CURB-65 score for treatment decisions in patients with CAP was associated with reduced mortality compared to the PSI score.	• Clinical outcomes according to the use of CURB-65 versus the PSI score should be studied in a randomized controlled trial.
Part 3: Treatment	
• We found no differences in clinical outcomes between patients with moderate-to-severe CAP treated orally or intravenously for based on the PSI score and clinical judgment.	• An RCT is warranted to compare oral versus intravenous treatment in moderate-to-severe CAP. Confirmation would justify guideline change.

Table 1. Summary of main findings and recommendations in this thesis. CAP = community-acquired pneumonia, PSI = pneumonia severity index, RCT = randomized controlled trial.

BROADER CONSIDERATIONS

Prediction models and artificial intelligence

In recent years, prediction models based on machine learning have become increasingly common. This raises the question whether machine learning approaches would provide advantages over the more ‘traditional’ regression models developed and evaluated in this thesis. When deciding on a modelling approach we emphasized future applicability, aiming for broadly applicable and explainable models that

would be easy to implement. It is known that machine learning has the potential to capture complex, non-linear relationships within large datasets, which may improve predictive performance compared to regression models. However, previous reviews suggest that when applied to relatively simple clinical data, the added value in predictive performance is often limited.^{50,51} Indeed, a machine learning bacteremia model showed similar predictive performance to our models.⁵² Furthermore, an important downside to machine learning-based models is that many function as ‘black boxes’, making it difficult to understand how a given prediction was generated. This lack of explainability might limit trust and willingness to use such models.⁵³ To address this, transparency of what predictors influence predictions is a minimum requirement. Explainable AI (XAI) has gained attention as a solution, with techniques like Shapley values increasingly being reported.⁵⁴ Finally, we hypothesize that models that are based on predictors with a clear biological rationale for their association with the outcome are likely to be more stable across a wider range of settings.¹⁹ Given these considerations, and our aim to develop robust and broadly applicable models based on non-complex clinical data, we chose to focus on a regression-based approach.

Aspect	Regression models	Machine learning models
Typical data	Simple, tabular data	Large datasets, possibly high-dimensional (images, free text)
Selection of predictors	Theory-based	Data-driven
Advantages	Transparent, easy to validate and apply in new settings, good explainability	Few assumptions needed, can capture complex, non-linear relationships and interactions
Disadvantages	Might underperform when true relations are complex or non-linear	Risk of overfitting, poor explainability due to ‘black box’ effect

Table 2. Comparison between more traditional regression models and machine learning models.

How to move forward to successful implementation?

Unlike much of contemporary medical research, this thesis does not focus on introducing additional interventions but rather asks the opposite question: what can safely be omitted? In the following section, several considerations that we believe are important for the successful implementation of this type of research are discussed.

First, it would be preferable if decisions regarding the use of prediction models to withhold tests or decide on treatment strategies, were not placed solely on individual clinicians. The potential negative consequences of such decisions are felt primarily

by the patient and their treating physician. For instance, consider a patient who is ultimately diagnosed with endocarditis after no blood cultures were obtained because of a low predicted risk of bacteremia. While our findings suggest that such cases would be rare, it is understandable that the physician involved would subsequently be hesitant to follow the model's recommendations in the future. Moreover, the benefits of reducing diagnostic testing such as lower costs and decreased resource utilization are primarily realized at the population level, leaving clinicians with little incentive to change their practice. Placing these decisions at a higher level, for instance through national expert groups such as the Dutch Working Party on Antibiotic Policy (SWAB), could help reduce such unwanted effects.

A further complication is that reasoning about risk itself is challenging. As discussed earlier, clinicians often strive to detect every single case of bacteremia, and the willingness to accept even a small risk of missing cases is low.¹⁶ Yet this creates an illusion of certainty. For example, studies have shown that bacteremia occurs in around 10% of patients with pyelonephritis treated as outpatients, without being associated with worse outcomes.⁵⁵ Even under current practice, therefore, some cases inevitably remain undetected. Therefore, we have to acknowledge that the clinical importance of identifying bacteremia varies by condition and patient context. In this regard, we could learn from primary care, where managing diagnostic uncertainty and accepting a certain degree of risk have long been more explicitly integrated in clinical decision-making.⁵⁶

Ultimately, a sense of collective responsibility is essential when considering the acceptance of potential increased risk, whether this is due to withholding diagnostic tests or by treating more patients with oral antibiotics. We should engage in broader discussions about what constitutes an acceptable level of risk for individual patients. Professional organizations such as the SWAB and the Dutch Internist Society could play a pivotal role by starting this dialogue and translating it into updated clinical guidelines.

CONCLUSIONS

This thesis suggests multiple improvements to the management of infectious diseases at the ED in three key areas: diagnostics, risk stratification and treatment. First, we developed prediction models that combine clinical variables with PCT. These models are promising to help avoid unnecessary blood cultures and contamination-related negative patient outcomes. Future validation and impact studies will help determine their clinical applicability. Second, we evaluated multiple tools for risk-stratification in a broad range of patients suspected of infection. This can help clinicians in deciding on site-of-care or treatment decisions in these specific

patient groups. Finally, we showed that there might be more room for oral antibiotics in the treatment of infections, potentially reducing admissions and therapy-related side-effects. Overall, this thesis represents a step forward to modern management strategies of infection by balancing safe reduction of testing and treatment without compromising safety.

REFERENCES

1. Jiménez-Aguilar P, López-Cortés LE, Rodríguez-Baño J. Impact of infectious diseases consultation on the outcome of patients with bacteraemia. *Ther Adv Infect Dis*. 2019 Dec 6;6:2049936119893576. doi: 10.1177/2049936119893576. PMID: 31839942; PMCID: PMC6900613.
2. Rhodes A, Evans LE, Alhazzani W et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Intensive Care Med*. 2017 Mar;43(3):304-377. doi: 10.1007/s00134-017-4683-6. Epub 2017 Jan 18. PMID: 28101605.
3. Lambregts MMC, Bernards AT, van der Beek MT, Visser LG, de Boer MG. Time to positivity of blood cultures supports early re-evaluation of empiric broad-spectrum antimicrobial therapy. *PLoS One*. 2019 Jan 2;14(1):e0208819. doi: 10.1371/journal.pone.0208819. PMID: 30601829; PMCID: PMC6314566.
4. Long B, Koyfman A. Best Clinical Practice: Blood Culture Utility in the Emergency Department. *J Emerg Med*. 2016 Nov;51(5):529-539. doi: 10.1016/j.jemermed.2016.07.003. Epub 2016 Sep 14. PMID: 27639424.
5. Coburn B, Morris AM, Tomlinson G, Detsky AS. Does this adult patient with suspected bacteremia require blood cultures? *JAMA*. 2012 Aug 1;308(5):502-11. doi: 10.1001/jama.2012.8262. Erratum in: *JAMA*. 2013 Jan 23;309(4):343. PMID: 22851117.
6. Nannan Panday RS, Wang S, van de Ven PM, Hekker TAM, Alam N, Nanayakkara PWB. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. *PLoS One*. 2019 Mar 21;14(3):e0214052. doi: 10.1371/journal.pone.0214052. PMID: 30897186; PMCID: PMC6428292.
7. Dempsey C, Skoglund E, Muldrew KL, Garey KW. Economic health care costs of blood culture contamination: A systematic review. *Am J Infect Control*. 2019 Aug;47(8):963-967. doi: 10.1016/j.ajic.2018.12.020. Epub 2019 Feb 20. PMID: 30795840.
8. Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: importance, implications, interpretation and prevention. *Clin Microbiol Infect*. 2018 Sep;24(9):964-969. doi: 10.1016/j.cmi.2018.03.030. Epub 2018 Apr 3. PMID: 29621616.
9. Bates DW, Goldman L, Lee TH. Contaminant blood cultures and resource utilization. The true consequences of false-positive results. *JAMA*. 1991 Jan 16;265(3):365-9. PMID: 1984535.
10. Kaal AG, Nieberg M, Stegmeijer K, Steyerberg EW, van Nieuwkoop C. The diagnostic accuracy of procalcitonin for community-acquired bacteremia; an updated systematic review and meta-analysis. [Submitted for publication]
11. Samsudin I, Vasikaran SD. Clinical Utility and Measurement of Procalcitonin. *Clin Biochem Rev*. 2017 Apr;38(2):59-68. PMID: 29332972; PMCID: PMC5759088.
12. Diamond IR, Grant RC, Feldman BM et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014 Apr;67(4):401-9. doi: 10.1016/j.jclinepi.2013.12.002. PMID: 24581294.
13. Fabre V, Hsu Y-J, Carroll KC et al. Multicenter evaluation of blood culture contamination and blood cultures practices in US acute care hospitals: time for standardization. *J Clin Microbiol*. 2025 Aug 13;63(8):e0053025. doi: 10.1128/jcm.00530-25. Epub 2025 Jul 11. PMID: 40643261; PMCID: PMC12345234.

14. Jones AE, Fiechtel JF, Brown MD, Ballew JJ, Kline JA. Procalcitonin test in the diagnosis of bacteremia: a meta-analysis. *Ann Emerg Med*. 2007 Jul;50(1):34-41. doi: 10.1016/j.annemergmed.2006.10.020. Epub 2006 Dec 11. PMID: 17161501.
15. Hoeboer SH, van der Geest PJ, Nieboer D, Groeneveld AB. The diagnostic accuracy of procalcitonin for bacteraemia: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2015 May;21(5):474-81. doi: 10.1016/j.cmi.2014.12.026. Epub 2015 Jan 14. PMID: 25726038.
16. Lautenbach E, Localio R, Nachamkin I. Clinicians required very high sensitivity of a bacteremia prediction rule. *J Clin Epidemiol*. 2004 Oct;57(10):1104-6. doi: 10.1016/j.jclinepi.2004.03.003. PMID: 15528062.
17. Kaal AG, Meziyerh S, van Burgel N et al. Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model. *J Infect*. 2024 Oct;89(4):106251. doi: 10.1016/j.jinf.2024.106251. Epub 2024 Aug 26. PMID: 39182652.
18. Kaal AG, Washer L, Wasylshyn A, Steyerberg EW, Mahajan P, van Nieuwkoop C. Bacteremia Prediction Models to Reduce Unnecessary Blood Cultures: External Validation in a Large US Emergency Department. [Submitted for publication]
19. Steyerberg EW. (2009) *Clinical prediction models: a practical approach to development, validation, and updating*. Springer, New York
20. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol*. 2022 Dec 12;22(1):316. doi: 10.1186/s12874-022-01801-8. PMID: 36510134; PMCID: PMC9742671.
21. Collins GS, de Groot JA, Dutton S et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014 Mar 19;14:40. doi: 10.1186/1471-2288-14-40. PMID: 24645774; PMCID: PMC3999945.
22. Buendía JA, Guerrero Patiño D. Cost-effectiveness of procalcitonin for detection of serious bacterial infections in children presenting with fever without source. *BMC Pediatr* 2022 Apr 26;22(1):226. <https://doi.org/10.1186/s12887-022-03293-3>. PMID: 35473509; PMCID: PMC9040337.
23. Voermans AM, Mewes JC, Broyles MR, Steuten LMG. Cost-effectiveness analysis of a procalcitonin-guided decision algorithm for antibiotic stewardship using real-world U.S. hospital data. *OMICS* 2019 Oct;23(10):508–15. <https://doi.org/10.1089/omi.2019.0113>. Epub 2019 Sep 11. PMID: 31509068; PMCID: PMC6806362.
24. Fabre V, Sharara SL, Salinas AB, Carroll KC, Desai S, Cosgrove SE. Does This Patient Need Blood Cultures? A Scoping Review of Indications for Blood Cultures in Adult Nonneutropenic Inpatients. *Clin Infect Dis*. 2020 Aug 22;71(5):1339-1347. doi: 10.1093/cid/ciaa039. PMID: 31942949.
25. WHO. *Global strategy on human resources for health: Workforce 2030*. 2016
26. Centers for Disease Control and Prevention Health Alert Network. Disruptions in availability of Becton Dickinson (BD) BACTEC™ blood culture bottles blood culture bottles. 2024. <https://emergency.cdc.gov/han/2024/han00512.asp>. Accessed 25th of June 2025.
27. Guly HR. Diagnostic errors in an accident and emergency department. *Emergency medicine journal*: 2001 Jul; 18(4): 263–269. doi: 10.1136/emj.18.4.263

28. Hussain F, Cooper A, Carson-Stevens A et al Diagnostic error in the emergency department: learning from national patient safety incident report analysis. *BMC emergency medicine* 2019 (1):77. doi:10.1186/s12873-019-0289-3
29. Jayaprakash N, Chae J, Sabov M, Samavedam S, Gajic O, Pickering BW. Improving Diagnostic Fidelity: An Approach to Standardizing the Process in Patients With Emerging Critical Illness. *Mayo Clinic proceedings Innovations, quality & outcomes* 3 2019 (3):327-334. doi:10.1016/j.mayocpiqo.2019.06.001
30. Dregmans E, Kaal AG, Meziyerh S et al. Analysis of Variation Between Diagnosis at Admission vs Discharge and Clinical Outcomes Among Adults With Possible Bacteremia. *JAMA Netw Open*. 2022 Jun 1;5(6):e2218172. doi: 10.1001/jamanetworkopen.2022.18172. PMID: 35737389; PMCID: PMC9226997.
31. van Daalen FV, Boersma WG, van de Garde EMW et al. Management of community-acquired pneumonia in adults: the 2024 practice guideline from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). 2024, SWAB. <https://swab.nl/nl/cap> Accessed on 26th of August, 2025.
32. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *New England Journal of Medicine*. 1997;336(4):243-250.
33. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003;58(5):377-382.
34. Chalmers JD, Singanayagam A, Akram AR, et al. Severity assessment tools for predicting mortality in hospitalised patients with community-acquired pneumonia. Systematic review and meta-analysis. *Thorax*. 2010;65:878–883. doi:10.1136/thx.2009.133280
35. Loke YK, Kwok CS, Niruban A, Myint PK. Value of severity scales in predicting mortality from community-acquired pneumonia: systematic review and meta-analysis. *Thorax*. 2010;65:884–890. doi:10.1136/thx.2009.134072
36. Ranzani OT, Prina E, Menéndez R, et al. New Sepsis Definition (Sepsis-3) and Community-acquired Pneumonia Mortality. A Validation and Clinical Decision-Making Study. *Am J Respir Crit Care Med*. 2017; 196: 1287–1297, Nov 15, 2017
37. Aujesky D, Auble TE, Yealy DM, et al. Prospective comparison of three validated prediction rules for prognosis in community-acquired pneumonia. *American Journal of Medicine*. 2005;118:384–392. doi:10.1016/j.amjmed.2005.01.006
38. Kaal AG, Op de Hoek L, Hochheimer DT et al. Outcomes of community-acquired pneumonia using the Pneumonia Severity Index versus the CURB-65 in routine practice of emergency departments. *ERJ Open Res*. 2023 May 2;9(3):00051-2023. doi: 10.1183/23120541.00051-2023. PMID: 37143846; PMCID: PMC10152258.
39. Huijts SM, Werkhoven CHv, Boersma WG, et al. Guideline adherence for empirical treatment of pneumonia and patient outcome. *The Netherlands Journal of Medicine*. 2013;71:502–507.
40. Chalmers JD, Rutherford J. Can we use severity assessment tools to increase outpatient management of community-acquired pneumonia? *Eur J Intern Med*. 2012;23:398–406. doi:10.1016/j.ejim.2011.10.002
41. Kaal A, Snel L, Dane M et al. Diagnostic yield of bacteriological tests and predictors of severe outcome in adult patients with COVID-19 presenting to the emergency department. *Emerg Med J*. 2021 Sep;38(9):685-691. doi: 10.1136/emered-2020-211027. Epub 2021 Jul 21. PMID: 34289966; PMCID: PMC8300553.

42. Parajuli P, Phipps LAC, Sabo R et al. The changing demographics and severity in hospitalized patients across COVID-19 variants: A national cohort study. *J Clin Transl Sci*. 2025 Jan 6;9(1):e22. doi: 10.1017/cts.2024.1166. PMID: 39911936; PMCID: PMC11795861.
43. Wunderink, R. G. Community-acquired pneumonia. *Infectious Diseases*. 2017:251–257.e1. doi: 10.1016/B978-0-7020-6285-8.00028-9. Epub 2016 Aug 12. PMCID: PMC7149822.
44. McLaughlin, J. M., Khan, F. L., Thoburn, E. A., Isturiz, R. E., Swerdlow, D. L. Rates of hospitalization for community-acquired pneumonia among US adults: A systematic review. *Vaccine*. 2020 Jan 22;38(4):741-751. doi: 10.1016/j.vaccine.2019.10.101. Epub 2019 Dec 13. PMID: 31843272.
45. Welte, T., Torres, A., Nathwani, D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax*. 2012 Jan;67(1):71-9. doi: 10.1136/thx.2009.129502. Epub 2010 Aug 20. PMID: 20729232.
46. Kaal AG, Roos R, de Jong P et al. Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study. *Sci Rep*. 2024 Apr 9;14(1):8271. doi: 10.1038/s41598-024-59026-2. PMID: 38594555; PMCID: PMC11004140.
47. Davar K, Clark D, Centor RM, et al. Can the Future of ID Escape the Inertial Dogma of Its Past? The Exemplars of Shorter Is Better and Oral Is the New IV. *Open Forum Infect Dis*. 2022;10(1):ofac706. Published 2022 Dec 29. doi:10.1093/ofid/ofac706. PMID: 36694838; PMCID: PMC9853939.
48. Wald-Dickler N, Holtom PD, Phillips MC, et al. Oral Is the New IV. Challenging Decades of Blood and Bone Infection Dogma: A Systematic Review. *Am J Med*. 2022;135(3):369-379. e1. doi:10.1016/j.amjmed.2021.10.007. PMID: 34715060; PMCID: PMC8901545.
49. Kaal AG, Roos R, Wilms EB, Borgdorff H, van Nieuwkoop C. Geef antibiotica bij voorkeur oraal [Choose oral over intravenous antibiotic therapy]. *Ned Tijdschr Geneesk*. 2024 Oct 15;168:D8300. Dutch. PMID: 39435708.
50. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019 Jun;110:12-22. doi: 10.1016/j.jclinepi.2019.02.004. Epub 2019 Feb 11. PMID: 30763612.
51. Nayebirad S, Hassanzadeh A, Vahdani AM et al. Comparison of machine learning models with conventional statistical methods for prediction of percutaneous coronary intervention outcomes: a systematic review and meta-analysis. *BMC Cardiovasc Disord*. 2025 Apr 23;25(1):310. doi: 10.1186/s12872-025-04746-0. PMID: 40269704; PMCID: PMC12016393.
52. Schinkel M, Boerman AW, Bennis FC et al. Diagnostic stewardship for blood cultures in the emergency department: A multicenter validation and prospective evaluation of a machine learning prediction tool. *EBioMedicine*. 2022 Aug;82:104176. doi: 10.1016/j.ebiom.2022.104176. Epub 2022 Jul 16. PMID: 35853298; PMCID: PMC9294655.
53. Elendu C, Amaechi DC, Elendu TC et al. Ethical implications of AI and robotics in healthcare: A review. *Medicine (Baltimore)*. 2023 Dec 15;102(50):e36671. doi: 10.1097/MD.00000000000036671. PMID: 38115340; PMCID: PMC10727550.

54. Allgaier J, Mulansky L, Draelos RL, Pryss R. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. *Artif Intell Med.* 2023 Sep;143:102616. doi: 10.1016/j.artmed.2023.102616. Epub 2023 Jun 24. PMID: 37673561.
55. van Nieuwkoop C, van't Wout JW, Spelt IC et al. Prospective cohort study of acute pyelonephritis in adults: safety of triage towards home based oral antimicrobial treatment. *J Infect.* 2010 Feb;60(2):114-21. doi: 10.1016/j.jinf.2009.11.008. Epub 2009 Nov 27. PMID: 19945482.
56. Alam R, Cheraghi-Sohi S, Panagioti M, Esmail A, Campbell S, Panagopoulou E. Managing diagnostic uncertainty in primary care: a systematic critical review. *BMC Fam Pract.* 2017 Aug 7;18(1):79. doi: 10.1186/s12875-017-0650-0. PMID: 28784088; PMCID: PMC5545872.



Appendices

NEDERLANDSE SAMENVATTING

Waar gaat dit proefschrift over?

Infecties zijn een van de meest voorkomende redenen waarom mensen naar de spoedeisende hulp (SEH) gaan. Denk aan longontsteking, een bloedvergiftiging, of een urineweginfectie. Als arts op de SEH moet je snel beslissen: welke onderzoeken doe ik? Hoe ernstig is de situatie? En welk antibioticum geef ik? Die beslissingen lijken misschien vanzelfsprekend, maar in de praktijk is dat lang niet altijd zo. Soms worden er meer onderzoeken gedaan dan nodig, of krijgen patiënten antibiotica via een infuus terwijl een pil net zo goed zou werken. Dat kost tijd, geld, en kan zelfs schadelijk zijn voor de patiënt.

Dit proefschrift gaat over de vraag: hoe kunnen we de behandeling van infecties op de SEH doelmatiger organiseren? Het doel is om de juiste zorg te geven aan de juiste patiënt. Niet te veel, maar ook zeker niet te weinig. Daarvoor onderscheiden we drie onderdelen: diagnostiek, risico-inschatting en behandeling.

Deel 1: Gerichter bloedkweken afnemen

Een van de meest gebruikte onderzoeken bij een vermoeden van infectie is de bloedkweek. Hierbij wordt bloed afgenomen en in het laboratorium gekweekt om te kijken of er bacteriën in het bloed zitten. Dat heet bacteriëmie, of in de volksmond 'bloedvergiftiging'. Het probleem is dat bloedkweken relatief duur zijn, en bovendien lang niet altijd iets opleveren: slechts 4 tot 11% van de kweken is daadwerkelijk positief. En soms is een 'positieve' uitslag helemaal geen bacteriëmie, maar een fout-positieve uitslag doordat er per ongeluk een bacterie van de huid in het flesje terecht is gekomen. Zo'n fout-positieve uitslag kan leiden tot onnodige opname, extra antibiotica, hogere kosten en in zeldzame gevallen zelfs meer kans op overlijden.

Wat hebben we onderzocht?

We hebben gekeken of we beter kunnen herkennen bij welke patiënten we veilig een bloedkweek achterwege kunnen laten. Hierbij hebben we gefocust op de bloedwaarde procalcitonine (PCT). PCT is een stof in het bloed die voornamelijk stijgt bij bacteriële infecties, en minder bij virale infecties. Uit een grote analyse van alle beschikbare studies bleek dat een lage PCT-waarde goed kan worden gebruikt om bloedvergiftiging uit te sluiten.

Om dit nog verder te verbeteren hebben we een model ontwikkeld dat een aantal metingen bij de patiënt, zoals leeftijd en bloeddruk, combineert met PCT. Dit model berekent per patiënt een geschat risico op een bacteriëmie. Bij een laag risico zou de arts dan kunnen beslissen de bloedkweek niet af te nemen.

Wat zijn de belangrijkste bevindingen?

We hebben dit model gebouwd en getest in Nederland, Spanje en Amerika. Het model bleek in staat om ongeveer 30% van de bloedkweken te besparen, terwijl slechts 1 tot 5% van de échte bacteriëmieën gemist wordt. Toekomstige studies zouden gericht moeten worden op het daadwerkelijk implementeren van het model in de praktijk.

Deel 2: Wie heeft een hoog risico op complicaties?

Het tweede deel van dit proefschrift gaat over het inschatten van risico: welke patiënten lopen de meeste kans op een ernstig ziektebeloop?

Verkeerde diagnose

We verwachtten dat patiënten met een verkeerde diagnose vaker een hoger risico op complicaties zouden hebben. Daarom hebben we onderzocht hoe vaak patiënten op de SEH een verkeerde infectiediagnose krijgen. Dat wil zeggen: de diagnose bij opname bleek achteraf niet te kloppen met wat er echt aan de hand was. Het resultaat was opvallend: bij 1 op de 9 patiënten met een vermoedelijke infectie klopte de diagnose niet. Gelukkig bleek dit in de meeste gevallen niet te leiden tot ernstige gevolgen zoals een langere ziekenhuisopname, opname op de intensive care, of overlijden. Wel kregen deze patiënten vaker brede antibiotica, waarschijnlijk omdat artsen twijfelden over de diagnose en het zekere voor het onzekere namen.

We hebben ook gekeken naar welke patiënten het vaakst een verkeerde diagnose kregen. Drie factoren sprongen eruit: hogere leeftijd, dementie, en een afwijkende urinetest zonder dat de patiënt klachten had die passen bij een urineweginfectie. Dat laatste is een bekende valkuil: ouderen hebben vaker bacteriën in de urine zonder dat ze ziek zijn, maar als de arts een afwijkende urinetest ziet, is de verleiding groot om toch een urineweginfectie te diagnosticeren.

Longontsteking: welke risicoscore werkt beter?

Voor patiënten met een longontsteking (pneumonie) zijn er twee veelgebruikte risicoscores om in te schatten hoe ernstig de situatie is: de PSI-score en de CURB-65. Artsen gebruiken deze scores als ondersteuning in de beslissing of iemand thuis kan blijven, opgenomen moet worden, of zelfs op de intensive care hoort. Ook wordt dit meegewogen in de beslissing welke antibiotica gegeven wordt.

In een landelijk onderzoek met bijna 51.000 patiënten hebben we vergeleken wat er gebeurt in ziekenhuizen die de CURB-65 gebruiken versus ziekenhuizen die de PSI score gebruiken. De uitkomst was verrassend: in ziekenhuizen die de CURB-65 gebruikten, overleden minder patiënten. Een mogelijke verklaring is dat de CURB-

65 vaker tot de conclusie 'ernstige pneumonie' leidt, waardoor patiënten vaker een zwaardere behandeling krijgen. Om beter te begrijpen of dit verschil echt een gevolg is van het gebruik van de scores, is meer onderzoek nodig. Maar mede gezien de betere gebruiksvriendelijkheid is het in de tussentijd te overwegen de CURB-65 boven de PSI score te verkiezen.

Deel 3: Tabletten in plaats van infuus

Het derde deel van dit proefschrift gaat over behandeling, en dan specifiek over de vraag: kunnen we meer patiënten met antibiotica in pilvorm behandelen (oraal), in plaats van een infuus? In de praktijk worden patiënten met een matig-ernstige longontsteking vaak opgenomen in het ziekenhuis en behandeld met antibiotica via een infuus. Dat vereist een ziekenhuisopname, is duurder, en brengt risico's met zich mee zoals bijvoorbeeld ontstekingen op de infuusplek. Maar als je kijkt naar hoe antibiotica worden opgenomen en afgebroken in het lichaam, zijn veel antibiotica even goed beschikbaar wanneer ze oraal worden ingenomen.

Wat hebben we onderzocht?

Het HagaZiekenhuis in Den Haag had een lokaal beleid waarbij artsen konden kiezen tussen orale antibiotica of een infuus. Dat week af van de nationale richtlijn, waarbij een infuus de aanbevolen standaard is. Dit gaf ons de kans om beide behandelgroepen te vergelijken.

Wat zijn de belangrijkste bevindingen?

We vonden geen verschil in uitkomsten tussen patiënten die orale antibiotica kregen en patiënten die een infuus kregen. Wel waren patiënten in de infuusgroep gemiddeld 2,5 dag langer opgenomen in het ziekenhuis. Dat suggereert dat orale behandeling niet alleen even veilig is, maar ook minder dagen in het ziekenhuis met bijbehorende kosten met zich meebrengt. Dit is een belangrijke bevinding, omdat het ingaat tegen de gangbare praktijk. Onze bevindingen bieden een basis voor een gerandomiseerde studie die dit definitief kan bevestigen.

Conclusie

Dit proefschrift laat zien dat er op meerdere vlakken ruimte is voor meer doelmatige zorg bij infecties op de spoedeisende hulp. Van gerichter afnemen van bloedkweken, beter herkennen welke patiënten risico lopen op complicaties, tot het vaker kiezen van orale behandeling boven een infuus. Dit proefschrift draagt bij aan een verbeterde aanpak van infectieziekten: met minder onnodige diagnostiek en behandeling, zonder in te leveren op veiligheid.

DANKWOORD

De afgelopen jaren hebben meer mensen een rol gespeeld in dit proefschrift dan ik op kan noemen, maar een paar wil ik hier graag specifiek bedanken.

Dit proefschrift zou er niet zijn zonder de geweldige steun van mijn co-promotor Kees van Nieuwkoop. Hij stond aan de wieg van het idee voor de PredictED studie en heeft mij de kans geboden om hier niet alleen aan mee te werken, maar om er een op maat gemaakt PhD traject van te maken. Bedankt voor de enorme hoeveelheid uren waarin we projecten hebben besproken, de steun voor al mijn ideeën en avonturen, en natuurlijk alle gezellige etentjes en borrels in Gouda. Ewout, ik had me geen betere promotor kunnen wensen. Ik heb niet alleen inhoudelijk heel veel van je geleerd, maar daarnaast heeft je scherpe gevoel voor wat daadwerkelijk werkt in de praktijk mij gevormd als onderzoeker. Dank voor je steun, en voor de waardering die ik altijd heb gevoeld.

Al het onderzoek gebeurt in een team, en er zijn een paar mensen die daarbij heel veel hebben geholpen. Lars, Pieter, Linde, Soufian en Emma, bedankt voor jullie enthousiasme, jullie creatieve ideeën, harde werk en gezelligheid binnen en buiten het ziekenhuis. Koen en Margot, dank voor jullie harde werk bij de oneindige stroom aan geïnccludeerde studies. Daarnaast was het heel bijzonder om met alle collega's van het ZorgInstituut te werken, waarbij we echt over de muren van onze organisaties hebben gekeken. Dank aan Davinia, Corline, Katrijn en Christel. Iemand die specifiek bereid was om samen te werken over de grenzen is Agustin. Door jou hebben we ons model in razend tempo buiten Nederland kunnen valideren. Dan alle collega's die mijn verblijf en project in Michigan niet alleen mogelijk hebben gemaakt, maar er ook een onvergetelijke levenservaring van hebben gemaakt: Savannah, Jackson, Monet, Brooke, Prashant, en ook Ingrid en Andrea.

Ik wil hier ook graag de ruimte nemen om de verschillende partijen te bedanken die dit onderzoek mogelijk hebben gemaakt: ten eerste de vakgroep Interne-Reumatologie van het HagaZiekenhuis, die mij in de gelegenheid hebben gesteld al mijn ideeën uit te werken en wiens steun ik altijd heb gevoeld. Verder het Cultuurfonds, Stichting Fulbright, het Leids Universitair Fonds, het Catharine van Tussenbroek Fonds en het Hendrik Mullerfonds.

Toen ik in januari 2023 plaatsnam in een van de kantoortjes op de afdeling Biomedical Data Sciences werd ik meteen opgenomen in het warme bad van alle promovendi. Dank voor jullie gezelligheid, voor alle leuke borrels en etentjes en natuurlijk jullie tijd en geduld als ik jullie hulp vroeg om mij de kneepjes van het vak predictiemodelleren bij te brengen. Dank in het bijzonder aan Ilaria, Frank, Ed, Jasper, Doranne, Lars en Marije.

Dit proefschrift is een weergave van onderzoek dat al jarenlang een constante aanwezigheid in mijn leven is geweest. Dat blijft natuurlijk niet onopgemerkt door je omgeving. Het is niet goed te beschrijven hoe waardevol het is om allemaal lieve familie en vrienden om je heen te hebben, en hoe groot de steun is die ik vanuit jullie allemaal heb ervaren. Een luisterend oor, het kritisch meedenken en meelesen, maar ook vooral het enthousiasme om het samen met mij te vieren als er weer een mijlpaal werd bereikt. Ik ben ontzettend dankbaar om jullie in mijn leven te hebben.

Lotte, wat was het leuk om naast jou te zitten en te zien hoe je van mijn verhalen een prachtig ontwerp maakt. Zonder jou had dit proefschrift er nooit zo mooi uitgezien.

Joost, als er iemand in mijn team zit dan ben jij het. Jij hebt alle pieken en dalen gezien en meegemaakt, en bent er altijd voor me geweest. Frustraties met R werden opgelost, je hielp me mee om het avontuur in Amerika te realiseren, en we hebben eindeloos gedanst in de woonkamer om de boel soms weer even helemaal op te schudden. Het leven zonder jou zou een vergissing zijn.

LIST OF PUBLICATIONS

1. Kaal A, Snel L, Dane M, van Burgel N et al. Diagnostic yield of bacteriological tests and predictors of severe outcome in adult patients with COVID-19 presenting to the emergency department. *Emerg Med J*. 2021 Sep;38(9):685-691. doi: 10.1136/emmermed-2020-211027. Epub 2021 Jul 21. PMID: 34289966; PMCID: PMC8300553.
2. Dregmans E, Kaal AG, Meziyerh S et al. Analysis of Variation Between Diagnosis at Admission vs Discharge and Clinical Outcomes Among Adults With Possible Bacteremia. *JAMA Netw Open*. 2022 Jun 1;5(6):e2218172. doi: 10.1001/jamanetworkopen.2022.18172. PMID: 35737389; PMCID: PMC9226997.
3. Kaal AG, Op de Hoek L, Hochheimer DT et al. Outcomes of community-acquired pneumonia using the Pneumonia Severity Index versus the CURB-65 in routine practice of emergency departments. *ERJ Open Res*. 2023 May 2;9(3):00051-2023. doi: 10.1183/23120541.00051-2023. PMID: 37143846; PMCID: PMC10152258.
4. Kaal AG, Roos R, de Jong P, Pepping RMC et al. Oral versus intravenous antibiotic treatment of moderate-to-severe community-acquired pneumonia: a propensity score matched study. *Sci Rep*. 2024 Apr 9;14(1):8271. doi: 10.1038/s41598-024-59026-2. PMID: 38594555; PMCID: PMC11004140.
5. Kaal AG, Meziyerh S, van Burgel N et al. Procalcitonin for safe reduction of unnecessary blood cultures in the emergency department: Development and validation of a prediction model. *J Infect*. 2024 Oct;89(4):106251. doi: 10.1016/j.jinf.2024.106251. Epub 2024 Aug 26. PMID: 39182652.
6. Kaal AG, Roos R, Wilms EB, Borgdorff H, van Nieuwkoop C. Geef antibiotica bij voorkeur oraal [Choose oral over intravenous antibiotic therapy]. *Ned Tijdschr Geneesk*. 2024 Oct 15;168:D8300. Dutch. PMID: 39435708.
7. Kaal AG, Nieberg M, Stegmeijer K, Steyerberg EW, van Nieuwkoop C. The diagnostic accuracy of procalcitonin for community-acquired bacteraemia: an updated systematic review and meta-analysis. *Clin Microbiol Infect*. 2026 May;32(5):749-759. doi: 10.1016/j.cmi.2025.12.029. Epub 2026 Jan 8. PMID: 41519266.
8. Kaal A, Washer L, Wasylshyn A, Steyerberg E, Mahajan P, van Nieuwkoop C. Bacteremia prediction models to reduce unnecessary blood cultures: external validation in a large US emergency department. *Sci Rep*. 2026 Apr 17;16(1):17928. doi: 10.1038/s41598-026-47024-5. PMID: 41998145; PMCID: PMC13249856.

KNOWLEDGE DISSEMINATION

Conferences

- 2021 Society to Improve Diagnosis in Medicine (SIDM): Poster presentation regarding misdiagnosis in the emergency department (Chapter 5)
- 2022 ID week: Poster presentation regarding CURB vs PSI for risk stratification in community-acquired pneumonia (Chapter 6)
- 2023 ESCMID: Poster presentation regarding oral vs intravenous treatment in community-acquired pneumonia (Chapter 8)
- 2023 ID week: Oral presentation regarding the PredictED study (Chapter 3)
- 2024 ID week: Poster presentation regarding external validation of bacteremia prediction models (Chapter 4)

Podcast

- 2024: Participation on podcast of NTVG to speak about oral versus intravenous use of antibiotics. Spotify episode 46: Kennis van de dood.

Other presentations

- 2022 Science day HagaZiekenhuis: Oral presentation regarding study in microbiological outcomes and severity prediction in COVID-19 (Chapter 7). Awarded with price for best presentation
- 2022 LUMC: Oral presentation regarding misdiagnosis in the emergency department for acute internal medicine fellows (Chapter 5)
- 2023 Science day HagaZiekenhuis: Two poster presentation regarding misdiagnosis in the emergency department (Chapter 5) and the use of CURB vs PSI for community acquired pneumonia (Chapter 6)
- 2024 Science day HagaZiekenhuis: Poster presentation regarding oral vs intravenous treatment in community-acquired pneumonia (Chapter 8)
- 2024 Science day HagaZiekenhuis: Oral presentation regarding the PredictED study (Chapter 3)
- 2025 A-team meeting and SWAB symposium: Oral treatment in moderate-severe community acquired pneumonia (Chapter 8)

Other publications

- Kaal AG, van Nieuwkoop C. Procalcitonine voor het reduceren van onnodige bloedkweken, EMPACT februari 2026. Interview about our research in reducing unnecessary blood cultures, published in a journal aimed to educate acute care health professionals.

CURRICULUM VITAE

Anna Kaal werd geboren op 28 november 1992 in Amsterdam. Nadat zij in 2003 naar Castricum verhuisde haalde zij haar VWO diploma aan het Jac. P. Thijssen College in Castricum. Na de middelbare school begon zij in 2010 aan de studie Technische Wiskunde aan de TU Delft, maar merkte al snel dat haar hart toch bij de zorg lag. Een jaar later begon zij alsnog aan de studie Geneeskunde aan de Universiteit van Leiden waarbij ze haar master een jaar lang combineerde met het Leiden Leadership Programme. Nadat zij de studie Geneeskunde cum laude had afgerond begon ze in 2018 als arts-assistent op de afdeling Interne geneeskunde in het HagaZiekenhuis in Den Haag. Daar kwam ze ook voor het eerst in aanraking met het idee voor de PredictED studie, waar ze aan ging werken onder begeleiding van Kees van Nieuwkoop, internist-infectioloog-acuut geneeskundige. Dit onderzoek legde de basis voor dit proefschrift. Vanaf mei 2020 begon zij aan de opleiding tot internist in het HagaZiekenhuis, wat zij combineerde met haar PhD. In deze jaren ontving zij meerdere beurzen die haar in de gelegenheid stelde om gedurende een halfjaar onderzoek te doen in de Verenigde Staten, waarna zij de opleiding vanaf september 2025 in het Leids Universitair Medisch Centrum vervolgde. Vanuit hier voltooide zij onder andere een stage infectieziekten in het Groote Schuur Ziekenhuis in Kaapstad, Zuid-Afrika. In het najaar van 2026 zal zij starten aan de differentiatie tot internist-infectioloog. Op dit moment woont ze samen met haar partner Joost in Den Haag.

